Identification and Characterization of a Novel Human Liver-specific Organic Anion Transporter (SLC22A7)

SIU SHU SHUN

A Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of Master of Philosophy

In

Biochemistry

Department of Biochemistry

Aug 2000

Acknowledgement

I would like to express my sincere thanks to Prof. K.P. Fung for giving me a valuable chance in studying this program. I am greatly indebted to Prof. C.Y. Lee, Prof. Mary M.Y. Waye and Dr. Stephen K.W. Tsui for their guidance and helpful suggestions in my research studies. I would also like to express my gratitude to colleagues of our laboratory, especially Enders Ng, Patrick Law, Simon Lee, Louis Kok and Masayo, their valuable comments and technical assistance support me in many aspects.

Contents

Abstract

From the screening of human liver cDNA library, we isolated a novel clone that has a high sequence identity (95%) to liver-specific transporter (NLT / hOAT2 / SLC22A7), and it is believed to be the human homolog of rat OAT2 / NLT. The 1647-nucleotide long hOAT2 cDNA encodes a protein with a deduced molecular weight of 60 kDa. It is regarded as the second member of organic anion transporter (OAT) because of its characteristic twelve putative transmembrane domains and sequence similarity to OAT1. HOAT2 possesses a large extracellular loop plus a possible *N*-glycosylation site within the loop.

It has been demonstrated that hOAT2 has a high level of expression in liver and relatively weak in kidney, while expression was only observed in liver when analyzed in fetal tissues, the mRNA transcript of hOAT2 was detected in cancer liver as well. Chromosomal localization of hOAT2 has been deduced to chromosome 6p21.2 – p21.1 by radiation hybrid mapping panel. Subcellular localization of hOAT2 using green fluorescence technique showed apparently to localize in cytosolic area of HepG2 and C2C12 cells. Anti-hOAT2 antibodies detected a band around ~60kDa from Hep3B total proteins.

摘要：

我們從人類肝臟的ｃＤＮＡ文庫可以分離到新的克隆，它的序列跟肝臟有關的載體９５％相同（ＮＬＴ／ｈＯＡＴ２／ＳＬＣ２２Ａ７），相信跟人類同族體。長１６４７個核甘酸的ｈＯＡＴ２的ｃＤＮＡ編成蛋白質的分子量.推斷爲６０ｋＤａ。由於它有獨特的十二條穿透細胞區段和序列跟ＯＡＴ１相似，它被分爲第二類有機陰離子載體（ＯＡＴ）。它擁有一條巨大的細胞外環在第一條和第二條穿透細胞區段之間，推斷會有一個Ｎ－醣基化的位置在環內出現。

ｈＯＡＴ２已被証實了在肝臟有大量的表現，但是相對地在腎臟較少，而只在胎兒的肝臟而非在腎臟表現·同時也在肝癌的ｈＯＡＴ２的ｍＲＮＡ抄本中檢測到。從放射混合繪畫技術ｈＯＡＴ２在染色體上的位置被推斷爲在染色體６ｐ２１.２至ｐ２１.１。而用綠色螢光技術表面顯示ｈＯＡＴ２在ＨｅｐＧ２和Ｃ２Ｃ１２的細胞質找到。利用反ｈＯＡＴ２抗體，發現Ｈｅｐ３Ｂ所有的蛋白質之中有一條６０ｋＤａ的紋。

Abbreviations

a.a.	Amino acid
BLAST	Basic Local Alignment Search Tool
bp	Base pair
cMOAT	Canalicular multispecific organic anion transporter
EDTA	Ethylenediaminetetracetic acid
ESTs	Expressed sequence tags
IPTG	Isopropylthio-b-D-galactoside
MOPS	(3-[N-Morpholino]propanesulfonic acid)
NCBI	National Center for Biotechnology Information
NLT	Novel liver-specific transporter
Ntcp	Na^+-dependent taurocholate cotransporting polypeptide
OAT	Organic anion transporter
OAT-K1	Kidney-specific oatp
Oatp	Organic anion transporter polypeptide
PCR	Polymerase chain reaction
PVDF	Polyvinylidene difluoride membrane
RT-PCR	Reverse transcriptase polymerase chain reaction
SDS	Sodium dodecyl sulfate
SLC22A7	Solute carrier family 22 member 7
TEMED	N,N,N',N'-Tetramethylethylenendiamine
X-gal	5-bromo-4-chloro-3-indolyl-b-D-galactoside

List of figures

List of tables

Chapter 1: Introduction

1.1 Human EST sequencing project, the role and goal

The rapid growth of international bioinformatics development and the combination with high throughput sequencing systems, a number of genetic sequences have been decoded in an exponential speed in the recent years. In 1995, the first bacterial genomic sequences have been announced (Fleischmann *et al.*, 1995). Yeast, worm and fly genomes have also been sequenced consecutively in 1996, 1998 and 1999 respectively. On June 26, 2000, the first draft sequence of the human genome has been completed, this announcement was made by the International Human Genome Project Consortium and Celera Genomics. Clearly, cracking the genetic code is a breakthrough to the scientific community, and certainly that such sequencing tasks are believed to take years to fully analyze.

Expressed sequence tags (ESTs) sequencing project is playing an another role, in addition to genomic sequence, in revealing the mystery of human genetic code. Large-scale partial sequencing of cDNA libraries to generate ESTs is widely utilized, which discovers novel genes and characterizing transcription profiles in various tissues, including the brain, the liver and lymphocytes (Adams

et al., 1991; Adams *et al.*, 1992; Adams *et al.*, 1993a; Adams *et al.*, 1993b; Okubo *et al.*, 1992).

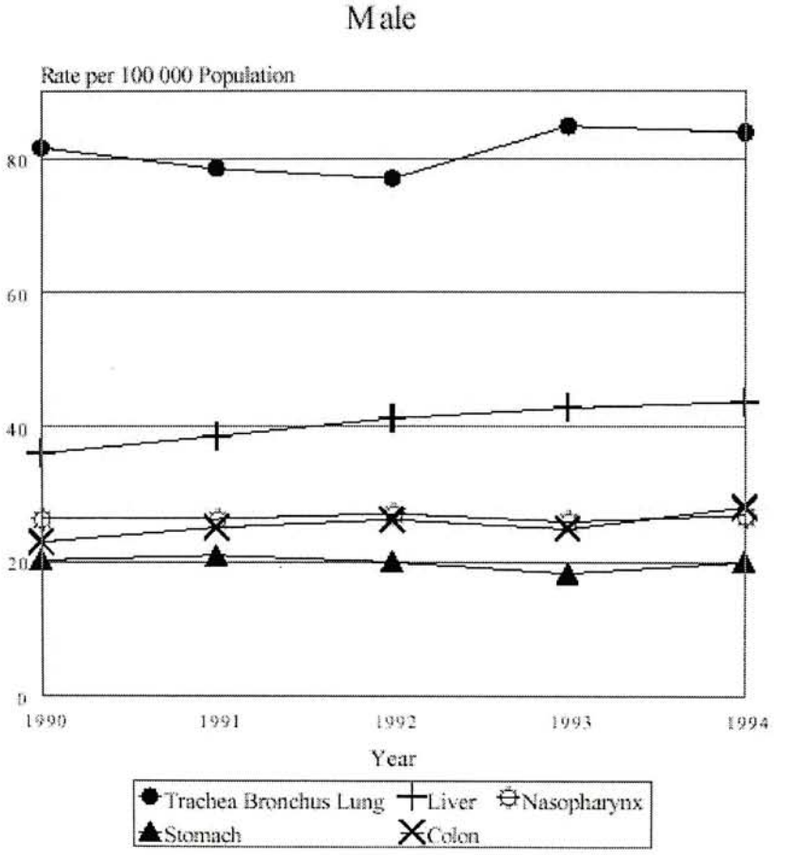
Data from single-pass sequencing of cDNA clones has also been used in other applications such as the construction of physical maps of chromosomes (Khan *et al.*, 1992; Wilcox *et al.*, 1991). Primers are designed according to the information from EST sequences, which are then used to define sequence-tagged sites (STSs) in the genome. STS is a short and operationally unique in the genome, which is used to generate mapping reagent after amplification with PCR (Olson *et al.*, 1989).

1.2 Human liver cDNA sequencing

Comparing with the western countries, it is not hard to observe that the incident rate of liver diseases such as cirrhosis, cholestasis and hepatitis is quite significant in Hong Kong. Statistical report 97/98 from Hospital Authority also figured out liver occupied the second high incident rate of leading cause of cancer in male (Figure 1.1). In our EST sequencing project, it mainly consists of three different sources of human liver cDNA libraries, normal liver from a Caucasian male, normal liver from a Chinese male and hepatitis induced cancer liver from an Chinese male. We target to analyze and differentiate the gene

expression patterns among the libraries, and to elucidate cancer induced and novel genes that are ubiquitously expressed and that may therefore be important in the general maintenance of cell function. In addition, we would like to develop EST sequence database that contains the information specific to hepatocellular system from Chinese people.

Figure 1.1 Incident rate of five leading causes of cancer in male from 1990 to 1994



1.3 The role of membrane-associated proteins in hepatocellular functions

1.3.1 Outline of the liver function

The liver is an excellent metabolic organ of vertebrates, which carries

various synthetic, metabolic functions and the detoxification of many endogenous and exogenous compounds such as excess amino acids, cholesterol, steroid hormones, bilirubin, toxins and many drugs (Frimmer and Ziegler, 1988; Meier, 1991). In other cases, it takes up various nutrients from the afferent blood, then it converts and stores them in carbohydrates, proteins and lipids. Upon nerve impulse or hormonal stimuli, these substrates will be released into the circulation. Whereas the movement of these substrates or ions in and out to a cell or organelle must pass through the basic structure, cell membrane.

1.3.2 Basic structure of hepatocyte

In the liver hepatocyte membrane structures, they include the nucleus, mitochondrion, peroxisome, endoplasmic reticulum, Golgi apparatus, lysosome, endosome, plasma membrane, and various types of vesicles. While the plasma membrane in the highly polarized hepatocyte is further differentiated into three separate membrane domains: apical domain, lateral domain and basal (sinusoidal) domain. The apical domain is a small microvillous area, which constitutes bile canaliculi and is involved in the secretion of bile salts into the bile. The lateral domain is a smooth surface bearing several types of junctions, mainly takes part in intercellular communications. The sinusoidal domain is facing the blood

circulation, which is participated in the secretion of serum proteins, uptake of molecules and communication with the biological system through its receptor that bind growth factors, hormones, and other biologically important ligands (Handler, 1989; Rodriguez-Boulan and Nelson, 1989). The distinct properties of these domains including transporting activities show the essential role playing by membrane proteins in which the proteins components largely confer the specificity of the membrane.

1.3.3 Category of membrane associated proteins

The distribution of human liver ESTs listed by functional categories is shown in the table 1.1. ESTs that are matched with known genes are classified into mainly eleven categories according to their apparent functions. It is noted that this large EST database is constructed by our cooperative works of our colleagues in our laboratory through the processes of cDNA sequencing, post-sequence editing, data analyzing, computer processing and other maintenance tasks. The proportions of membrane-associated proteins only accommodate a low percentage but roughly similar among the cDNA libraries for normal (3.5%), normal-tumor (3.6%) and tumor (2.0%) tissues.

Table 1.1 Frequency distribution of human liver true normal, tumor-normal and tumor ESTs classified by functional categories.

Category	Normal lib	Tumor-normal lib	Tumor lib
	Frequency (%)	Frequency (%)	Frequency (%)
Contractile element	2 (0.4)	6 (0.2)	8 (0.2)
Cytoskeletal related	6 (1.1)	35 (1.2)	60 (1.4)
Extracellular matrix	0 (0.0)	11 (0.4)	34 (0.8)
Energy metabolism	14 (2.5)	116 (3.9)	254 (6.1)
Hormone and hormonal control	7 (1.2)	33 (1.1)	35 (0.8)
Signal transduction/cell cycle regulation	17 (3.0)	80 (2.7)	201 (4.8)
Transcription and translation	21 (3.7)	473 (15.9)	1264 (30.3)
Membrane associated	19 (3.3)	108 (3.6)	83 (2.0)
Secretory protein	335 (58.5)	1281 (43.1)	764 (18.3)
Other metabolism	93 (16.3)	437 (14.7)	857 (20.6)
Miscellaneous	58 (10.2)	393 (13.2)	613 (14.7)
Subtotal	570 (100)	2973 (100)	4173 (100)
Novel	147	804	1218
Mitochondrial related	83	845	673
Grand Total	800	4622	6064

1.4 Identification of human OAT2 gene

For over 800 cDNA partial sequences from human liver EST project, about one-fifth of them were identified as novel category. The first notification of OAT2 gene was found in the novel gene category. A cDNA clone called aa402, which is closely matched (84% nucleotide identity) with *Rattus norvegicus* liver-specific transport protein (NLT) (Simonson *et al.*, 1994).

The first isolation of rat organic anion transporter 1 (OAT1) was performed by Sekine *et al.* (1997). Physiological and pharmacological studies and sequence alignment between OAT1 and NLT, revealed that they are similar in both multispecificity, characteristic tissue distribution and sequence identity (Sekine *et al.*, 1998). To this point, NLT was supposed to be a member of the OAT family and be renamed as OAT2.

Kim, RB *et al.* (Unpublished) submitted the human liver-specific transporter (NLT) mRNA sequence to the GenBank database in 1999. Although its molecular function or transportation property is not yet characterized, the high nucleotide sequence similarity between rat OAT2 and human NLT made us to speculate that human NLT is the human homolog of rat OAT2.

For standardizing purpose, the human OAT2 gene was sent to HUGO Gene Nomenclature Committee (HGNC) for assigning a formal gene name. Solute

carrier family 22 member 7 (SLC22A7) was denoted as alternative name for hOAT2.

1.5 The multispecific transporter family

1.5.1 Classification

When lipophilic xenobiotics are transported into hepatocytes and extensively biotransformed by P450- and transferase-mediated reactions. The produced negatively charged amphiphilic compounds must be secreted from the hepatocytes. As a result, multispecific transporters of anionic compounds are necessary for the liver (Meier *et al.*, 1995; Petzinger E *et al.*, 1994).

Up till now, at least three multispecific transporter families were identified, they are ATP-binding cassette (ABC) superfamily, organic anion transporting polypeptide (oatp) superfamily and organic anion transporter (OAT) superfamily. The ATP-binding cassette (ABC) superfamily includes P-glycoprotein, a multidrug resistance associated protein and a canalicular multispecific organic anion transporter (cMOAT) which is considered to mediate hepatobiliary excretion of organic anions. Na^+-independent oatp, prostaglandin transporter and kidney-specific oatp (OAT-K1) are belongs to oatp superfamily. The OAT superfamily was newly discovered when the distinct structural and transport

characteristics of OAT1 had been identified (Sekine *et al.*, 1997). To date, four members of OAT include OAT1, OAT2, OAT3 and OAT4 have been sequenced and characterized.

1.5.2 The human OAT family

The striking feature of this organic anion transport system is its wide substrate selectivity, covering not only endogenous substrates and also a number of important therapeutic drugs (Cihlar *et al.*, 1999). Are there any associations among these transporters, in molecular basis or physiological reaction? Is there any common binding domain for recognizing organic anions? Studies on structurally related isoforms could facilitate our understanding on their multispecific properties.

hOAT1 is found to strongly express in kidney and weakly in brain (Hosoyamada, 1999). Numerous organic anions including endogenous compounds, xenobiotics and their metabolites are eliminated by kidney, which are extracted though the basolateral membrane. Physiological studies show it mediates probenicid- and alpha-ketogultarate-sensitive *para*-aminohippurate (PAH) transport when expressed in *Xenopus Laevis* oocytes (Race *et al.*, 1999). Another two, possibly a result of alternatively spliced sequences of hOAT1, the

hOAT1-1 and hOAT1-2, were isolated by Hosoyamada *et al.* (1999). Whereas the difference between hOAT1-2 and hOAT1-1 is the 13-amino acid deletion located near the carboxyl terminal on hOAT1-2. In addition to the mediation of PAH transport by hOAT1-1, it was also demonstrated that hOAT1-1 is inorganic anions dependent and the replacement of extracellular Cl^- with other anions or mannitol suppressed the PAH uptake (Hosoyamada *et al.*, 1999). This phenomenon has been reported other than *Xenopus Laevis* oocytes expression system such as sliced rabbit kidney (Goldinger *et al.*, 1980), basolateral membrane vesicles from bovine kidney (Schmitt and Burckhardt, 1993) and basolateral membrane vesicles from rat kidneys (Inui *et al.*, 1986).

hOAT3 shows 46% amino acid sequence identity to hOAT1. Chromosomal mapping found that both of the hOAT1 and hOAT3 genes were localized to chromosome 11. Northern blot analysis demonstrated they were highly expressed in kidney (Race *et al.*, 1999). The apparent similarities do not confer related functional properties in the oocytes expression experiments. Organic anions such as PAH, urate and TEA that were found to be increased uptake by hOAT1 did not appear to be the substrates of hOAT3.

The human OAT4 was the latest member among the OAT family, which was identified by Cha *et al.* (2000) in a human kidney cDNA library screening.

Although the crystalline structures of the four transport proteins have not yet been elucidated, they share a common feature of twelve-putative transmembrane domains. The multispecific property was also elucidated in hOAT4. It mediated the transport of estrone sulfate, dehydroepiandrosterone sulfate (DHEA-s) and orcharatoxin A, where DHEA-s is a precursor of estrogen. The choice of these substrates was in consideration of the tissue specificity of hOAT4 protein and high expressions were detected in placenta and kidney (Cha *et al.*, 2000). Whereas the uptake of DHEA-s by the placenta is important in synthesizing estrogens (estrone, estradiol, and estriol), and also keep the fetus away from the cytotoxicity of DHEA-s (Cha *et al.*, 2000).

1.6 The characteristics of rat multispecific OAT2

To date, the functional or molecular studies on the human OAT2 protein have not yet been clearly defined. One could provide some clues as to the understanding of hOAT2 is its rat homolog, rOAT2 (NLT). ROAT2 gene had been sequenced by Simonson *et al.* (1994). It encodes a 535 amino acid protein that possesses twelve putative transmembrane regions. A preferential expression of rOAT2 was detected in liver and rather weak in kidney by Northern blot analysis (Sekine *et al.*, 1998). Indirect immunostaining of rat liver section was

used to investigate the rOAT2 location in hepatocyte. Monoclonal antibody 2C5, which was specific to the transporter, detected the highest signal on the sinusoidal domain, rather than the lateral and the canalicular regions. Furthermore, the expression of the rat OAT2 is developmentally regulated. In liver development, it showed a very low level in 19-21 gestation period to nearly a 5-fold increase in post-partum stage (Simonson *et al.*, 1994).

Considerations on the functional characterization of rOAT2 protein were based on two main categories, the organic anions multispecificity and the liver-specific properties. A group of organic anions including α-ketoglutarate, prostaglandin E_2, acetylsalicylate, PAH and therapeutic drug, methrotrexate were sorted in the uptake experiments using the oocytes expression system (Sekine *et al.*, 1998). ROAT2 had been demonstrated significant mediation of these anions, except for the low-level uptake of PAH was detected. It was one of the evident to show the similar structure of rOAT1 to rOAT2 did not represent having similar transport efficacy, where a much higher level of PAH transport was detected in rOAT1 (Sekine *et al.*, 1997). Table 1.2 summarizes the various characteristics of OATs from human and rat.

Table 1.2 Molecular and functional characteristics of organic anion transporters in this study.

Designation	Species	Amino acids	Tissue distribution	Substrate	Accession	Reference
OAT1	Rat	551	Kidney	PAH, glutarate	AB004559	(Sekine *et al.*, 1997)
OAT2 (NLT)	Rat	535	Liver, Kidney	Salicylate, α-ketoglutarate, PGE_2	L27651	(Simonson *et al.*, 1994)
OAT3	Rat	536	Liver, Kidney, Brain	PAH, ochratoxin A, estrone sulpfate, cimetidine	AB017446	(Kusuhara *et al.*, 1999)
OAT1-1	Human	563	Kidney	PAH	AB009697	(Hosoyamada *et al.*, 1999)
SLC22A6, (OAT1), hOAT1-2	Human	550	Kidney	PAH	AF124373	(Cihlar *et al.*, 1999; Hosoyamada *et al.*, 1999; Race *et al.*, 1999)
SLC22A7 (OAT2)	Human	548	Liver, Kidney	Unknown	AF097518	This study, (Kim *et al.*, unpublished)
OAT3	Human	568	Kidney	Unknown	AF097491	(Race *et al.*, 1999)
OAT4	Human	550	Kidney, Placenta	Estrone sulfate, DHEA-s, ochratoxin A, PAH	AB026116	(Cha *et al.*, 2000)

PAH, *p*-aminohippurate; PGE2, prostaglandin E_2; DHEA-s, dehydroepiandrosterone sulfate; NLT, novel liver-specific transporter.

1.7 Clinical significance of organic anion transport proteins

Cholestasis is one of the focuses in studying hepatic membrane transporters. Either the impairment of bile secretion from the hepatocyes into the canaliculi or the obstruction to the flow of preformed bile through the biliary tract will possibly leading to this disease. Bile acid secretion from sinusoidal blood into bile is mediated by several kinds of transport proteins locating in the sinusoidal (basolateral) hepatocye membrane.

Na^+-taurocholate cotransporting polypeptide (Ntcp) is one of the involved transporters responsible for bile acid uptake in a sodium-dependent manner (Hagenbuch and Meier, 1994; Hagenbuch *et al.*, 1991). Based upon experimental studies on rats, Ntcp showed a downregulation of taurocholate uptake in cholestasis such as bile duct ligation (Gartung *et al.*, 1996) and endotoxinemia (Bolder *et al.*, 1997; Green *et al.*, 1996). This finding was further demonstrated in human cholestatic liver with primary biliary cirrhosis. A significant inverse relationship between Ntcp mRNA levels and extrahepatic biliary atresia and plasma bilirubin concentration indicated downregulation of Ntcp was also found in man (Shneider *et al.*, 1997). Such inhibitory regulation of bile acid uptake probably serves to protect the hepatocyte from further intracellular accumulation of potentially toxic bile salts in case of the canalicular efflux systems for bile

acids are impaired (Bossard *et al.*, 1993).

In addition to Ntcp, members of sodium-independent organic anion transporting polypeptides (Oatp), Oatp1 and Oatp2, were also found to be down-regulated in ethinyl estradiol cholestasis (Simon *et al.*, 1996) and during liver regeneration (Gerloff *et al.*, 1999). Such differential regulation of these basolateral organic anions transporters in regenerating liver might protect replicating liver cells by diminishing uptake of potentially hepatotoxic bile salts, because the remaining liver initially cannot cope with the original bile acid pool size (Gerloff *et al.*, 1999)

Numerous drugs and endogenous compounds are excreted via a transport system, organic anion transporters such as OAT1 and OAT3 were seemed to be related in mediating nephrotoxicants like β-lactam antibiotics, mycotoxins and sevoflurane degradation product (compound A) (Endou, 1998). For example, mycotoxins such as ochratoxin A (OTA), which was clarified as a rate-limiting step in nephrotoxicity through the transmembrane transport by OAT1 (Jung and Endou, 1989).

Chapter 2: Materials and Methods

2.1 Human liver EST sequencing project

2.1.1 Plating out the adult human liver phage library

A single colony of *Escherichia coli* bacterial host (XL1-Blue MRF') was inoculated in 50ml of LB broth supplemented with 0.2% (w/v) maltose and 10mM Mg_2SO_4, with shaking at 30°C in 250rpm overnight. The cells were centrifuged at 2,000rpm (Beckman rotor JLA-16.250) at 4°C for 10min and the cell pellet was resuspended in 15ml of 10mM Mg_2SO_4.

100 to 150 pfu of the λ-ZAP phage of adult human liver cDNA library (Clontech) was added to 200μl of host cells in 10mM Mg_2SO_4. The phage bacteria mixture was then incubated at 37°C for 15min. Concomitantly, 2ml of NZY top agar prewarmed to 48°C with 15μl of 0.5M IPTG and 10μl of X-gal (50mg/ml) were prepared. The incubated mixture plus the top agar solution were then plated onto prewarmed at (37°C) NZY plates and allowed to set for 10min, and then the plates are incubated at 37°C overnight.

The semi-transparent plaques were picked and transferred to 50μl of suspension medium buffer (100mM NaCl, 9mM Mg_2SO_4, 50mM Tris-HCl, 0.01% gelatin) which allowed the phage to diffuse. The phage stocks were ready for PCR

and stored at 4°C.

2.1.2 PCR detection and amplification of the cDNA clone

PCR was performed using forward and reverse primers of λ-ZAP (5'-CCAAGC TCGAAA TTAACC CTC –3') and M13 Universal primer (5'-CGACGT TGTAAA ACGACG GCCAGT –3') respectively. The 50μl reaction mixture included 2μl phage suspension, 2.5mM each of the forward and reverse primers, 5μl 10X reaction buffer (Pharmacia), 0.5μl *Taq* polymerase (2.5U/μl, Pharmacia), and 5μl 2mM dNTP (Pharmacia). The reaction started by 30 cycles of denaturing at 95°C for 30sec, followed by 50°C for 35sec and 72°C for 2 min. The final extension proceeded at 72° for 5min using a programmable thermal cycler PTC-100 (MJ Research, Inc.).

After the cycling reactions were completed, the PCR products were analyzed on a 1% TBE (0.1M Tris-HCl, 0.08M boric acid, 1mM EDTA) agarose gel for checking insert sizes and purity. In order to obtain an efficient sequencing productivity, this step is crucial to isolate vectors that have very small inserts, no insert or disability to be amplified. The PCR product that produced a single obvious band under the electrophoresis was chosen for further DNA sequencing.

2.1.3 Automatic cDNA sequencing

Approximate 150ng of PCR product was used as template, with additional to 2μl 10X sequencing buffer, 1U Thermosequenase DNA polymerase and 5pmol fluorescein conjugated CY5-T3 sequencing primer (5'-GCTCGA AATTAA CCCTCA CTAAAG -3') to work as a master mix. They were then added separately into 4μl termination mix dATP/ddATP, dTTP/ddTTP, dCTP/ddCTP and dGTP/ddGTP. A drop of mineral oil was added on the surface of the liquid layer to prevent vaporization during the reaction. The sequencing reaction was started at 94°C for 35sec, 50°C for 35sec and 72°C for 1min 30sec. 30 cycles were repeated and a final extension at 72°C for 10min. Afterwards, 5μl of stop solution (95% formamide, 20mM EDTA and 10mg/ml blue dextran) was mixed to terminate the reaction.

The reaction mixtures individually terminated with the four dideoxynucleotides ddATP, ddTTP, ddCTP and ddGTP were electrophoresed on a 6% polyacrylamide gel by using automated DNA sequencer (ALF automated sequencer, Pharmacia™) for 6 hours.

2.2 Cloning of hOAT2 gene into TA cloning vector pT-Adv

2.2.1 Amplification of hOAT2 by PCR

One μl of cDNA from human fetal liver (Invitrogen) was used as templates for a total of 50μl PCR mixture. Forward (5'- ATGGGC TTTGAG GAGCTG CTGGA) and reverse primers (5'- TTAGTT CTGGAC CTGCTT CATGGG) which flank the two ends of hOAT2 gene with 0.2mM dNTPs, 1X PCR buffer and 1U DNA polymerase mix from Clontech Advantage® cDNA polymerase kit were added to commence the reaction. PCR was performed according to the following protocol: 94°C for 30 s; 36 cycles of 94°C for 30 s, 60°C for 45 s and 68°C for 2 min; 68°C for 5 min. The reaction conditions were manipulated by PTC-200 thermal cycler (MJ Research Inc.).

2.2.2 Ligation reaction

A 10μl volume of ligation mixture composed of approximately 25ng of freshly prepared PCR product, 1μl of 10X ligation buffer (Clontech), 2μl of 25ng/μl pT-Adv vector (Appendix), 1μl T4 DNA ligase (4.0 Weiss units/μl, Clontech) and sterile water. Finally, it was taken to incubation at 16°C overnight.

2.2.3 Transformation of recombinant plasmid into competent cells

Before transformation, one 50-μl of frozen TOP10F' *E.coli* competent cells (Clontech) were thawed on ice. 2μl of ligation product along with 2μl of β-mercaptoethanol were pipetted into the tube and incubated on ice for 30min. Then it was heat shocked in the 42°C water bath for 30sec and immediately transferred into ice for 2 min. The cells were recovered by adding 250μl of SOC medium, shaked at 225rpm at 37°C for 1hr. 10μl of X-Gal (50mg/ml) and IPTG (400mM) with 100μl transformation mixture were spreaded on LB plate containing 50μg/ml of either kanamycin or ampicillin at 37°C overnight. Colonies in white were ready to pick for further analysis.

2.3 Sequence analysis and structural prediction

The BLAST is a set of similarity program (Altschul *et al.*, 1994) containing large nucleotide and peptide sequences database which provides a fast and updated sequence similarity search. The ESTs and other genetic sequences similarity searches were determined by BLAST, National Center for Biotechnology Information, Bethesda, Maryland (NCBI) (http://www.ncbi.nlm.nih.gov/BLAST). Amino acid and nucleotide sequence alignments were performed using BLAST 2 sequences (NCBI) for pair alignments (Tatusova and Madden, 1999), ClustalW

(freely downloaded from http://www.sgi.com/chembio/resources/clustalw/parallel_clustalw.html) and DNASIS for multiple alignments. Prediction of transmembrane regions and orientation in protein sequences was done by DNASIS and HMMTOP. While the transmembrane helices and topology for transmembrane proteins were analyzed using Hidden Markov Model (Tusnady and Simon, 1998). The presence of motifs in the amino acid sequence hOAT1 was predicted using PROSITE scan from Network Protein Sequence Analysis (NPS@) at Pôle Bio-Informatique Lyonnais (PBIL), France (http://pbil.ibcp.fr/NPSA/npsa_proscan.html).

2.4 Cloning of the hOAT2 gene into the pQE30 expression vector

2.4.1 PCR amplification and restriction endonuclease cutting

Restriction cutting sites *Bam*HI and *Sal*I were designed on the forward and reverse primers respectively to amplify the hOAT2 gene:

Forward: 5'- TAGGGC GGATCC GGCTTT GAGGAG CTGCTG GAGCAG -3';

Reverse: 5'- TAGGGC GTCGAC TTAGTT CTGGAC CTGCTT CATGGG -3', whereas end-clamps consist of TAGGGC flanking the 5' region of both primers were aimed to facilitate the restriction enzymes digestion

PCR was performed according to the following conditions: 94°C for 30 s; 36 cycles of 94°C for 30 s, 60°C for 45 s and 72°C for 2 min; 72°C for 5 min. Restriction cutting of the PCR product was prepared in a total volume of 20μl composed of 2X One-Phor-All PLUS buffer (Pharmacia), 0.5μl of each restriction enzymes, *Bam*HI and *Sal*I, and approximately 1μg of PCR product. Then the reaction was carried out at 37°C for three hours as well as the restriction digestion was done for the expression vector pQE30 (Appendix) by those enzymes in the same conditions.

2.4.2 Gene clean

The restriction digested PCR product and vector were purified by using GENECLEAN™ III kit (Bio 101 Inc.) in order to remove excess salt and the unwanted digested fragments. The DNA was electrophoresed on a 1% TAE agarose gel for about 20min in constant voltage of 120V. The desired band was excised from the gel and placed in 3 volumes of 6M sodium iodide. It was then incubated in 55°C water bath for 10min until the gel was completely solublized. 10 to 15 μl of GLASSMILK which depends on the DNA amount was mixed and incubated at room temperature for 15 min with occasional shaking. They were centrifuged at 15,000 x g for 30 sec. The supernatant was discarded, and the pellet

was washed with 500µl of ice-cold NEW WASH solution three times. The pellet was dried by speed vacuum before the same volume of sterile water to GLASSMILK was added to resuspend the pellet and eluted the bound DNA. Afterwards, the DNA extracted after centrifugation was collected and subsequently quantified by running on a 1% TBE agarose gel.

2.4.3 Preparation of bacterial competent cells

A single bacterial colony of *Escherichia coli* M15(pREP4) was inoculated with 2ml LBK (25µg/ml kanamycin) medium at 37°C overnight with shaking at 250rpm. 250µl of the overnight culture was inoculated into 25ml LB medium containing the appropriate antibiotics which was then incubated at 37°C for about 3 hours with shaking at 250rpm until it reached log phase. The culture was collected in a 50 ml centrifuge tube.

The bacterial cells were spun down by centrifuging at 3,000rpm (Beckman rotor JLA-16.250) for 15 min at 4°C. The pellet was resuspended in one-third volume of original culture of RF1 buffer (100mM RbCl, 50mM $MnCl_2 \cdot 4H_2O$, 30mM potassium acetate, 10mM $CaCl_2 \cdot 2H_2O$, 15% w/v glycerol, pH 5.8). The bacterial cells were incubated on ice for 15 to 20 min before centrifuging at 3,000 rpm (Beckman rotor JLA-16.250) at 4°C for 15 min. The pellet was then resuspended in

1/12.5 volume of the original culture of RF2 buffer (10mM MOPS, 10mM RbCl, 75mM $CaCl_2 \cdot 2H_2O$, 15% w/v glycerol, pH 6.8). The bacterial cells were chilled on ice for another 15 to 20 min. Thereafter, they were aliquoted into 200µl fractions contained by pre-chilled microcentrifuge tubes for later use. The competent bacterial cells were frozen in liquid nitrogen and stored at -70°C.

2.5 Small scale synthesis of plasmid DNA

The mini-preparation of plasmid DNA was performed by using Promega Wizard® plus SV Minipreps DNA Purification System. After transforming the recombinant plasmid DNA in appropriate *E.coli* strain and plated out, a single colony was picked and inoculated in a 2ml culture and incubated at 37°C with shaking overnight. The cell culture was then poured into a 1.5ml microcentrifuge tube and centrifuged at 10,000 x g for 5 min. After the supernatant was removed, 250µl of Cell Resuspension Solution was added with vortexing. Afterwards, 250µl of Cell Lysis solution, 10µl of Alkaline Protease Solution and 350µl of Neutralization Solution were mixed accordingly by inverting the tube and incubated not longer than 5 minutes between each solution was added. The bacterial lysate was centrifuged at 14,000 x g for 10min, the supernatant was transferred to a Minipreps Spin Column. Under centrifugation at 14,000 x g for a

minute, plasmid DNA trapped in the column was further washed with Column Wash Solution and at last the DNA was eluted with 60μl of sterile water.

2.6 Large scale synthesis of the plasmid DNA

A single colony was picked from master plate and inoculated in a starter 2ml culture for 8 hours at 37°C, 250 rpm. 10ml of log phase culture was diluted in 100ml selective medium for another growth at 37°C for overnight, 250rpm.

The cells were centrifuged at 6,000 x g at 4°C for 15min and the pellet were resuspened in 10ml of buffer P1 (100μg/ml Rnase A, 50mM Tris-HCl, 10mM EDTA pH 8.0). Cells were lysed 10ml of buffer P2 (1% SDS, 200mM NaOH), and the tube was inverted several times and incubated for 5min at room temperature. Another 10ml of chilled buffer P3(20mM KOAc, pH 5.5) was mixed to the lysate, then the lysate were immediately poured into the QIAfilter Cartridge (QIAGEN) and filtered after 10 minutes incubation.

Purification of plasmid DNA was achieved by running the filtered lysate into QIAGEN anion-exchange resin column which was pre-equilibrated by 10ml buffer QBT (750mM NaCl, 50mM MOPS, pH 7.0, 15% isopropanol). Majority of contaminants from the plasmid preparation were removed by washing the column with 30 ml buffer QC (1.0M NaCl, 50mM MOPS, pH 7.0, 15% isopropanol) twice.

Afterwards, 15ml of buffer QF (1.2 M NaCl, 50mM MOPS, pH 8.5, 15% isopropanol) was added for elution. DNA collected was precipitated by 10.5ml isopropanol and centrifuged at 15,000 x g for 30min at 4°C. The pellet was washed with 5 ml 70% ethanol and further centrifuged at 15,000 x g for 10min. Finally, the purified plasmid DNA was redissolved by adding on appropriate volume of sterile water.

2.7 Cloning of the hOAT2 gene into the pSecTag2B mammalian expression vector

In additional to the end clamp sequence TAGGGC, restriction sites *BamHI* in forward primers and *XhoI* site in reverse primers were placed in the hOAT2 cDNA by flanking its ends.

Forward primer 5'- TAGGGC GGATCC GGGTTT GAGGAG CTGCTG GAGCAG -3'

Reverse primer 5'- TAGGGC CTCGAG TGTTCT GGACCT GCTTCA TGGGCA T -3'

The conditions used in the PCR amplification were described in cloning of pQE30 vector. After a series of processes such as restriction cutting, ligation and transformation into bacterial host for vector propagation, the transformants were plated onto LBA selective plates (100μg/ml).

2.8 Cloning of the hOAT2 gene into the pEGFP-C1 fluorescent vector

Amplification of the hOAT2 inserts were performed using the forward primer (5'- AGGGCC TCGAGC TATGGG CTTTGA GGAGCT GCTGGA -3') and the reverse primer (5'- TAGGGC GGATCC TTAGTT CTGGAC CTGCTT CATGGG -3') which synthesized the transcripts flanking with two restriction sites *XhoI* and *SalI* respectively. An end-clamp, TAGGGC, was added to the 5' end of the primers to enhance the restriction site cutting. The PCR was done according to the profile that was mentioned above with little modification of the annealing temperature to 62°C for 45 s.

After restriction enzymes cutting, they were cloned into the green fluorescent vector pEGFP-C1. The transformants were plated onto LBK plates with 25μg/ml of kanamycin.

2.8.1 Tissue culture and transfection

C2C12 myoblasts, was maintained in RPMI-1640 medium or DMEM medium (Life Technologies, Inc.) supplemented with 10% (vol/vol) heat-inactivated, certified Fetal Bovine Serum (Life Technologies Inc.), 1% antibiotics (penicillin, 100 units/ml; streptomycin sulfate, 100 μg/ml) at 37°C in 5% CO_2 in a humidified atmosphere. Cells were subcultured every 2 days.

The recombinant expression vectors carrying the insert hOAT2 were purified

as described in this chapter. C2C12 myoblasts was transfected with various plasmid constructs using Lipofectamine Plus Reagent (Life Technologies, Inc.). Approximate 1×10^5 Cells were grown on coverslips in six-well plates, until the cells were 60-70% confluent. 1 μg of DNA and 4 μl of the Plus Reagent were added to 125 μl RPMI-1640 medium lacking serum and antibiotics and incubated for 15 min at room temperature. 4 μl of lipofectamine reagent were added to 125 μl medium without serum and antibiotics to another new tube and incubated for 15 min at room temperature. After incubation, these two solutions were mixed and incubated for another 15 min at room temperature. The cells were washed with PBS twice. 1 ml of serum free medium was added to the cells. The DNA-lipofectamine plus mixture was added to the cells and incubated for 3 hours at 37°C. After that, the cells were rinsed with PBS and fresh medium supplemented with 10% FBS and antibiotics was added followed by incubation for another 24 hours before microscopic examination.

2.8.2 Fluorescence microscopy examination

For epi-fluorescence microscopy (Zeiss Axioskop fluorescence microscopy), cells expressing GFP fusion protein was viewed with either a GFP filter set (Omega Optical, Inc.). Images were captured by cooled CCD camera

(Photometeric, Inc) (exposure 1 to 4 s) and were artificially colored using IpLab spectrum software.

2.9 Chromosomal mapping of the hOAT2 gene

2.9.1 Somatic cell hybrids mapping

Gene specific primers were designed from the coding sequence of the hOAT2 gene (Forward: 5'- GGCTGG GAGTAC GACCAC TCAGAA T; Reverse: 5'- AGGCCC AGCACC AGGGTA CTCAC). PCR was first performed to a 24 somatic cell hybrids DNA panel (human / rodent somatic cell hybrids mapping panel 2-Version 3; National Institute of General Medicine Service Coriell Cell Repositories) (Drwinga *et al.*, 1993) using the Advantage polymerase kit (Clontech). The working conditions used were as follows: initial denaturation at 94°C for 1 min; 32 cycles of 94°C for 30 s, 65°C for 30 s, 68°C for 1.5 min; and a final extension at 68°C for 5 min. The presence of gene specific bands was analyzed on 2% agarose / EtBr gel.

2.9.2 Radiation hybrids mapping

A higher fidelity map was applied by a PCR-based radiation hybrid mapping technique using the Genebridge 4 human / hamster whole-genome Radiation

Hybrid Panel RH02.05 (Gyapay *et al.*, 1996). It includes 93 hybrid genomic DNAs from the same number of human-on-hamster somatic cell lines, and also two control genomic DNAs (HFL human, A23 hamster). Twenty-five ng of genomic DNA was used for the PCR using the same primers performed in somatic hybrids mapping. The final products were checked on 2% agarose gel. The resulting 1-kb PCR product was subcloned into the TA cloning vector (Clontech) and sequenced for confirmation.

The scores of "1" and "0" were used to represent the presence and absence of specific bands respectively, while ambiguous bands were taken as "2". Duplicate typings were performed before analysis. A string of scores generated from the screening results were processed by the software RHMAPPER at the Whitehead Institute / MIT Center for Genomic Research (Lee *et al.*, 1998).

2.10Reverse Transcription Polymerase Chain Reaction (RT-PCR)

The steps of mRNA preparation and the conversion from mRNA to cDNA by reverse transcriptase were avoided by using Multiple Tissue cDNA (MTC) Panels (Clontech). They are sets of normalized, first-strand cDNA generated using RNA from different human tissues. A set of PCRs were performed whereas the primers set and condition profile were optimized according to various MTC panels.

Forward primer F1 (5'- ATGGGC TTTGAG GAGCTG CTGGA) and reverse primer R2 (5'- ATCTCT TCCTCC TGAAGA CTGGTT) were chosen in human normal MTC panel. For human cancer and fetal MTC panels, forward primer F2 (5'- GATCTG GTGTGT GAGCAG AAAGG) was utilized instead of F1. One-fifth ng of each cDNA preparation was amplified in a total volume of 25μl. The same PCR profile was used in all panels of first denaturing at 94°C for 1 min; 35 cycles of 94°C for 30 s, 57°C for 45 s, 72°C for 2 min; and final extension at 72°C for 10 min. The PCR products were analyzed in 1% agarose gel.

The relative abundance of specific transcripts within panels was normalized using a housekeeping gene glyceraldehyde-3-phosphate dehydrogenase (G3PDH). 10μM of G3PDH Control Primer Mix provided in the kit with upstream primer: 5'- TGAAGG TCGGAG TCAACG GATTTG GT -3' and downstream primer: 5'- CATGTG GGCCAT GAGGTC CACCAC -3' that amplify the G3PDH gene in a ~1-kb fragment.

2.11 Western hybridization

2.11.1 Preparation of anti-hOAT2 antibodies

2.11.1.1 Synthetic peptide conjugation

A peptide sequence which is specific to human consisting 15 amino acids (NTTLGEERQSRGELE, postion 91-105) at the amino-terminus of human OAT2 was synthesized. 0.1mg of the synthetic peptide was conjugated with keyhole limpet hemocyanin (KLH) solution (5mg/ml) for 30 min. NHS (5mg/ml) and EDAC (10mg/ml) were added with stirring for 30 min. This conjugate was then purified with a G-25 Sephadex desalting column (Pharmacia) equilibrated with PBS. The cloudy eluent was aliquoted and stored at -20°C, and ready for injection.

2.11.1.2 Immunizing rabbit polyclonal antibodies for human OAT2

0.5ml of the KHL-hOAT2 peptide (100μg) was emulsified by mixing 0.5ml of Freund's complete adjuvent for the first subcutaneous injection into a rabbit. Incomplete Freund's adjuvent was used in the boosts six weeks after priming immunization. Ten to forteen days after the boost, the first blood was collected. Boost immunization and blood collection were continued in 4-week interval, the rabbit was killed at the fifth blood drawing. The blood was clotted at room temperatre before it was centrifuged at 5,000 x g at 4°C to remove blood clot and

cell debris. Finally the serum was stored at -20°C.

2.11.1.3 Purification of the rabbit polyclonal IgG antibodies

Purification of the rabbit anti-hOAT2 antibodies was done by running through a HiTrap protein G affinity column (Pharmacia). The column was prepared by removing the top-cap and a drop of start buffer (0.02 M sodium phosphate, pH 7) was applied to the top of the column to avoid air bubbles. The ethanol preservative was washed out from the column with 3ml start buffer. The column was equilibrated with 3 bed volumes of start buffer. 2ml of serum separated from rabbit's blood was applied onto the column using a syringe. The column was washed with 5ml start buffer. Then, the column was eluted with 4ml elution buffer (0.1M glycine-HCl, pH 2.7) and collected in microcentrifuge tubes of each with 1ml fraction. Each 100 µl neutralizing buffer (1.0M Tris-HCl, pH 9.0) was added to each fraction to restore the pH around 7. The purified polyclonal IgG was aliquoted and stored at -20°C.

2.11.2 Western blot analysis

2.11.2.1 Protein isolation from rat liver

Liver tissue resected from rat was homogenized in TRIZOL Reagent

(GibcoBRL) in the concentration of 75mg/ml. Three phases were separated after adding chloroform and followed by centrifugation. The lower red, phenol-chloroform phase was aspirated, 0.3ml of 100% ethanol per 1ml of TRIZOL Reagent used for the initial homogenization was mixed. After sedimentation, the phenol-ethanol supernatant was isolated for protein preparation. 1.5ml of isopropanol per 1 ml of TRIZOL Reagent was added, the protein was centrifuged at 12,000 x g for 10min at 4°C. The protein pellet was washed three times with washing solution (0.3M GuHCl in 95% ethanol) and centrifuged at 7,500 x g for 5 min at each wash. The cleaned protein pellet was dissolved in 1% SDS solution and stored at -20°C.

2.11.2.2 Protein preparation from cell lysate

A 20mm flask containing about 1x 10^6 cells were ready for protein collection. The medium in the culture flask was discareded and then the cells were washed with PBS three times. A scraper was used to detach the cells from the flask and a trace amount of PBS was used to resuspend the cells. The cells were moved to a 1.5 ml eppendorf and centrifuged at 1,500 x g for 5 min. 50-100 µl of lysis buffer mixed with protease inhibitor was added to resusupend the pellet, and incubated on ice for 15 min. Incubation in room temperature or higher is avoided to minimize

protein degradation. Finally, proteins were collected from the supernatant after further centrifugation at 10,000 x g for 10 min at 4°C.

2.11.2.3 Quantitation of total proteins by Bradford protein assay

Protein concentration was determined using Bio-Rad protein assay kit II. Protein standard was set up by diluting the standard Bovine Serum Albumin (BSA) solution into concentration of 0.9, 0.6, 0.3 and 0.2 mg/ml. The BioRad's Coomassie Reagent was five-fold diluted with deionized water while the protein samples were divided into 1/5, 1/10 and 1/20 dilutions. To each 20µl of protein sample, 1ml of Coomassie Reagent was mixed and inverted 4 times. The sample mixture was allowed to incubate at room temperature for 5min for color development, absorbance at 595nm was taken thereafter.

2.11.2.4 Blotting and hybridization

Proteins from various sources were denatured by mixing loading buffer (2% SDS, 125mM Tris-HCL, pH 7.4, 20% glycerol, 2% β-mercaptoethanol) at 100°C for 5 minutes and electrophoresed on a 12% SDS-polyacrylamide gel thereafter. The gel was soaked in transfer buffer (25mM Tris, 192 mM glycine, 20% v/v methanol, pH 8.3) for 15 minutes before the proteins were transferred to

polyvinylidene difluoride (PVDF) membranes (Millipore) by Mini Trans-Blot electrophoretic transfer cell at 200V for 2 hours.

The blots were blocked with 5% non-fat milk in Tris buffered saline (TBS)-T (20mM Tris, 137mM sodium chloride, 1M hydrochloric acid, 0.05% Tween-20, pH 7.6) overnight at 4°C and incubated with the affinity-purified anti-hOAT2 antibodies (1:300 dilution) for 1 h at room temperature. Goat anti-rabbit IgG conjugated with horseradish peroxidase (HRP, 1:3000 dilution, Amersham Pharmacia Biotech) secondary antibodies were used to incubate with the blots for another hour at room temperature. Enhanced chemiluminescence kit was applied for detection (Amersham Pharmacia Biotech), and the signal was exposed onto BioMax MR films (Kodak). For identification of the antibody specificity, the anti-hOAT2 antibody was pre-incubated with 10μg of the immunogen peptide.

Chapter 3: Results

3.1 Catalogue of the 500 liver ESTs

The distribution of the 500 liver ESTs after homology search in BLAST was presented in the Table 3.1. Within the ESTs, 300 of them were generated from human normal liver cDNA library and the remaining 200 ESTs were generated from cancerous liver, as they were matched in according to their major functions. The overview of the gene expression pattern of these two libraries was summarized in the Table 3.2 below.

Table 3.2 The overall pattern of gene expression in the 500 liver ESTs.

Category	Tumor liver		Normal liver	
	Frequency	%	Frequency	%
Matched genes				
Contractile element (CE/CO)	0	0.0%	0	0.0%
Cytoskeletal related (CY)	4	3.0%	2	0.9%
Extracellular matrix (EX)	1	0.7%	0	0.0%
Energy metabolism (EN)	6	4.5%	4	1.8%
Hormone and hormonal control (H)	1	0.7%	2	0.9%
Signal transduction/cell cycle regulation (CS)	10	7.5%	4	1.8%
Transcription and translation (TT)	42	31.3%	7	3.1%
Membrane associated (M)	4	3.0%	5	2.2%
Secretory protein (OS)	32	23.9%	146	65.2%
Other metabolism (OM)	21	15.7%	37	16.5%
Miscellaneous (MI)	13	9.7%	17	7.6%
Subtotal	134	100.0%	224	100.0%
Overall				
Matched genes	134	75%	224	80%
Novel	26	14%	24	8%
Mitochondrial	20	11%	33	12%
Total	180	100%	281	100%

Table 3.1 The human liver ESTs that were matched to known genes, either human or non-human matches, were classified into eleven categories. Each category as shown in the table includes the accession number, distribution frequency and *P*-value. a) The 281 EST sequences of $P < 0.001$ generated from 300 clones of normal adult liver cDNA library.

Subject	Accession#	P Value	Class
casein kinase 1 alpha (c)	gi|1432066	e-114	CS
cell cycle regulatory protein p95 (NBS1)	gi|3098674	e-165	CS
COP9 complex subunit 7a (COPS7a) (b)	gi|3309173	5e-83	CS
Gephyrin (a)	gi|56311	e-111	CS
myosin light chain kinase isoform 3B (MLCK)	gi|4322023	e-166	CY
tubulin, beta-, clone 24464	gi|3387979	e-159	CY
electron transfer flavoprotein beta subunit	gi|297901	0	EN
glucokinase regulator	gi|683571	1e-35	EN
isocitrate dehydrogenase, NADP-dependent (IDH)	gi|3641397	e-140	EN
succinate dehydrogenase flavoprotein subunit (SDH)	gi|347133	e-100	EN
angiogenin	gi|178249	e-126	H
angiotensinogen	gi|178643	e-156	H
gp96, human homologue of murine tumor rejection antigen	gi|37260	e-111	M
HLA-DMA (HLA-DMA34 allele)	gi|450815	e-103	M
low-affinity IgG Fc receptor (beta-Fc-gamma-RII)	gi|184843	8e-87	M
major histocompatibility complex (HLA-B*44031)	gi|1929863	5e-69	M
TGN46 protein	gi|1518268	3e-83	M
autonomously replicating sequence (ARS)	gi|179295	0	MI
haptoglobin related (Hpr)	gi|32446	1e-69	MI
haptoglobin related protein	gi|1495457	e-109	MI
HIV-1 TAR RNA binding protein (TARBP-b)	gi|347963	e-129	MI
KIAA0108	gi|285962	0	MI
KIAA0214 protein	gi|3869186	0	MI
Ku (p70/p80) subunit	gi|186793	7e-68	MI
metallothionein-IG (MT1G), (clone 14VS)	gi|188712	5e-58	MI
mRNA clone with similarity to L-glycerol-3-phosphate:NAD oxidoreductase and albumin gene sequences.	gi|763428	0	MI
polyA site DNA	gi|505034	0	MI
polyA site DNA	gi|505035	9e-75	MI
Proline synthetase associated	gi|4126977	0	MI
RasGAP-related protein (IQGAP2)	gi|1262925	7e-62	MI
SDP3	gi|3126978	2e-49	MI
selenoprotein P	gi|36425	0	MI
unknown protein	gi|393314	0	MI
unknown protein	gi|393314	e-106	MI
isolate Afr7 cytochrome b (CYTB) gene, mitochondrial gene encoding mitochondrial protein	gi|3372408	0	MITO
isolate Afr7 cytochrome b (CYTB) gene, mitochondrial gene encoding mitochondrial protein	gi|3372378	e-123	MITO
isolate Afr7 cytochrome b (CYTB) gene, mitochondrial gene encoding mitochondrial protein	gi|3372376	0	MITO
mitochondrial DNA, complete genome	gi|1262342	e-141	MITO
mitochondrial DNA, complete genome	gi|1262342	e-139	MITO
mitochondrial DNA, complete genome	gi|1262342	e-125	MITO
mitochondrial DNA, complete sequence	gi|644480	0	MITO
mitochondrial DNA, complete sequence	gi|644480	0	MITO
mitochondrial genome	gi|13003	e-172	MITO
mitochondrial genome	gi|13003	e-124	MITO
mitochondrial genome	gi|13003	e-148	MITO
mitochondrial genome (consensus sequence)	gi|13000	e-119	MITO

Subject	Accession#	P Value	Class
mitochondrial genome (consensus sequence)	gi|13000	5e-98	MITO
mitochondrial genome (consensus sequence)	gi|13000	1e-97	MITO
mitochondrial genome (consensus sequence)	gi|13000	0	MITO
mitochondrial genome (consensus sequence)	gi|13000	e-132	MITO
mitochondrial genome (consensus sequence)	gi|13000	e-162	MITO
mitochondrial genome (consensus sequence)	gi|13000	e-138	MITO
mitochondrial genome (consensus sequence)	gi|13000	0	MITO
mitochondrial genome (consensus sequence)	gi|13000	e-132	MITO
mitochondrial genome (consensus sequence)	gi|13000	2e-74	MITO
mitochondrion, complete genome	gi|1944628	e-154	MITO
mitochondrion, complete genome	gi|1944628	4e-79	MITO
mitochondrion, complete genome	gi|1944628	e-129	MITO
mitochondrion, complete genome	gi|1944628	e-178	MITO
mitochondrion, complete genome	gi|1944628	0	MITO
mitochondrion, complete genome	gi|1944628	e-137	MITO
mitochondrion, complete genome	gi|1944628	e-151	MITO
mitochondrion, complete genome	gi|1944628	0	MITO
mitochondrion, complete genome	gi|1944628	2e-96	MITO
mitochondrion, complete genome	gi|1944628	1e-50	MITO
NADH dehydrogenase subunit 2 (ND2) gene, mitochondrial gene encoding mitochondrial protein	gi|3287312	0	MITO
NADH dehydrogenase subunit 2 (ND2) gene, mitochondrial gene encoding mitochondrial protein	gi|3287320	0	MITO
Apolipoprotein H (beta-2-glycoprotein I)	gi|28811	e-109	N
BAC clone GS165I04 from 7q21	gi|2275190	2e-91	N
BAC clone RG066D11 from 7q22	gi|2335066	8e-45	N
BAC clone RG161K23 from 7q21	gi|1809224	0	N
Chromosome 16 BAC clone CIT987SK-582J2	gi|3219333	2e-84	N
Chromosome 16 BAC clone CIT987-SKA-113A6 ~complete genomic sequence	gi|2815548	0	N
Chromosome 16 BAC clone CIT987SK-A-259H10	gi|3337388	e-165	N
Chromosome 16 BAC clone CIT987SK-A-259H10	gi|3337388	e-119	N
Chromosome 16 BAC clone CIT987SK-A-923A4	gi|3219338	1e-67	N
chromosome 16, cosmid clone 363E3 (LANL)	gi|3097841	9e-31	N
cosmid clone LUCA13 from 3p21.3	gi|2337875	e-106	N
DNA sequence from clone 390O13 on chromosome Xp11	gi|3204451	0	N
DNA sequence from clone 425C14 on chromosome 6q22	gi|3281967	3e-90	N
DNA sequence from clone 662N3 on chromosome Xp21.1-21.3.	gi|4455593	4e-13	N
DNA sequence from cosmid 65O19, BRCA2 gene region chromosome 13q12-13 contains dinucleotide repeat	gi|1322149	0	N
DNA sequence from PAC 408N23 on chromosome 22q13	gi|2582746	5e-60	N
glycoprotein, alpha-1 acid	gi|178256	0	N
Human immunodeficiency virus type 1 isolate F6E envelope glycoprotein (env)	gi|733351	0	N
Xp22 BAC GSHB-184P14 (Genome Systems Human BAC library)	gi|3273378	6e-70	N
adenylate kinase 1 (AK1), nuclear gene encoding mitochondrial protein	gi|4502010	e-147	OM
alcohol dehydrogenase beta-1-subunit (ADH1-2 allele)	gi|28415	e-157	OM
aldehyde dehydrogenase 1 (ALDH1)	gi|2183298	0	OM
aldehyde reductase	gi|178480	8e-97	OM
aldolase B	gi|28616	e-110	OM
aldolase B	gi|178352	e-143	OM
aldolase B	gi|28616	e-105	OM
arginase, liver	gi|178994	e-161	OM
argininosuccinate synthetase	gi|28871	e-110	OM
C1 inhibitor	gi|29536	0	OM

Subject	Accession#	P Value	Class
C1 inhibitor	gi|29536	2e-87	OM
carbamyl phosphate synthetase I	gi|3228247	e-127	OM
carbamyl phosphate synthetase I (EC 6.3.4.16)	gi|219552	4e-17	OM
cysteine dioxygenase type 1	gi|467560	8e-83	OM
cytochrome P-3-450	gi|30338	8e-66	OM
cytochrome P-3-450	gi|181307	e-163	OM
cytochrome P450 2E1 (CYP2E1)	gi|4249652	0	OM
cytochrome P-450j	gi|181359	0	OM
fatty acid binding protein, liver (FABP)	gi|182355	e-159	OM
ferritin L chain	gi|182513	e-159	OM
ferritin L chain	gi|182513	e-147	OM
ferritin L-chain	gi|2230868	e-136	OM
ferritin light subunit	gi|182517	2e-65	OM
flavin-containing monooxygenase form II (FMO2)	gi|188630	e-100	OM
homogentisate 1,2-dioxygenase	gi|1561615	e-164	OM
I-FLICE isoform 5	gi|2827297	e-149	OM
KDEL receptor, a presumptive	gi|34030	0	OM
lambda-crystallin (c)	gi|164904	7e-48	OM
lanosterol 14-demethylase cytochrome P450 (CYP51)	gi|1698395	e-117	OM
nifedipine oxidase	gi|181373	0	OM
oxidative 3 alpha hydroxysteroid dehydrogenase	gi|2661210	5e-79	OM
phosphoglucomutase 1 (PGM1)	gi|189925	e-138	OM
serin protease with IGF-binding motif	gi|1513058	e-168	OM
serine dehydratase	gi|338029	e-121	OM
serine-pyruvate aminotransferase	gi|36581	e-109	OM
tyrosine aminotransferase (TAT) (EC 2.6.1.5)	gi|36712	0	OM
uridine diphosphoglucose pyrophosphorylase	gi|881393	e-106	OM
acid glycoprotein, alpha-1	gi|178256	e-144	OS
acid glycoprotein, alpha-1	gi|178256	0	OS
acid glycoprotein, alpha1- (orosomucoid)	gi|24444	0	OS
acid glycoprotein, alpha1- (orosomucoid)	gi|24444	e-118	OS
acid glycoprotein, alpha1- (orosomucoid)	gi|24444	0	OS
acid glycoprotein, alpha1- (orosomucoid)	gi|24444	6e-96	OS
acid glycoprotein, alpha1- (orosomucoid)	gi|24444	e-102	OS
albumin (ALB)	gi|4502026	3e-54	OS
albumin (ALB)	gi|4502026	0	OS
albumin (ALB)	gi|4502026	0	OS
albumin (ALB)	gi|4502026	e-177	OS
albumin (ALB)	gi|4502026	0	OS
albumin (ALB)	gi|4502026	1e-77	OS
albumin (ALB)	gi|4502026	2e-81	OS
albumin (ALB)	gi|4502026	e-171	OS
albumin (ALB)	gi|4502026	0	OS
albumin (ALB)	gi|4502026	0	OS
albumin (ALB)	gi|4502026	e-151	OS
albumin (ALB) mRNA	gi|4502026	e-121	OS
albumin (ALB) mRNA	gi|4502026	2e-96	OS
alpha-1-microglobulin/bikunin precursor (AMBP)	gi|4502066	e-134	OS
antichymotrypsin, alpha 1- gene fragment	gi|28331	e-121	OS
antichymotrypsin, alpha1-	gi|439142	e-153	OS
antichymotrypsin, alpha-1-, precursor	gi|4165889	0	OS
antithrombin III (AT3)	gi|4502260	e-114	OS
antitrypsin, alpha 1-	gi|28965	e-179	OS
antitrypsin, alpha 1-	gi|28965	0	OS
antitrypsin, alpha 1-	gi|28965	e-103	OS
antitrypsin, alpha-1-	gi|177828	1e-07	OS
antitrypsin, alpha-1-	gi|177828	e-159	OS
antitrypsin, alpha-1-	gi|28965	e-160	OS
antitrypsin, alpha-1-	gi|28965	1e-32	OS

Subject	Accession#	P Value	Class
antitrypsin, alpha-1-	gi|28965	e-151	OS
antitrypsin, alpha-1-	gi|177828	e-165	OS
antitrypsin, alpha-1-	gi|28965	1e-94	OS
antitrypsin, alpha-1-	gi|177828	e-156	OS
antitrypsin, alpha-1-	gi|177828	0	OS
antitrypsin, alpha-1-	gi|177828	0	OS
antitrypsin, alpha-1- (S variant)	gi|177830	4e-88	OS
antitrypsin, alpha-1- (S variant)	gi|177830	e-155	OS
antitrypsin, Z type alpha-1-	gi|177835	2e-90	OS
apolipoprotein B100	gi|178789	e-168	OS
apolipoprotein B-100	gi|178803	2e-80	OS
apolipoprotein B-100	gi|178817	e-111	OS
apolipoprotein C-III (apoC-III)	gi|178828	e-154	OS
apolipoprotein J	gi|178854	0	OS
carboxypeptidase B, prepro-plasma	gi|189686	e-151	OS
complement component 4B (C4B)	gi|4502500	e-168	OS
complement component C3 (C3)	gi|1162925	0	OS
complement component C3 (C3)	gi|1162925	e-150	OS
complement component C3 , alpha and beta subunits	gi|179664	e-159	OS
complement component C3, alpha and beta subunits	gi|179664	e-106	OS
complement component C3, alpha and beta subunits	gi|179664	e-153	OS
complement component C3, alpha and beta subunits	gi|179664	e-100	OS
complement component C3, alpha and beta subunits	gi|179664	e-128	OS
complement component C9	gi|29580	2e-87	OS
complement component C9	gi|29580	e-127	OS
complement cytolysis inhibitor (CLI)	gi|180619	e-147	OS
complement factor B	gi|297568	e-146	OS
complement factor B	gi|297568	e-138	OS
complement factor B	gi|297568	e-148	OS
complement factor H	gi|31964	0	OS
C-reactive protein	gi|181067	5e-97	OS
C-reactive protein	gi|30223	e-162	OS
C-reactive protein	gi|30212	2e-99	OS
fibrinogen A-alpha-chain	gi|182425	0	OS
fibrinogen A-alpha-chain	gi|182425	e-153	OS
fibrinogen alpha chain	gi|458553	e-134	OS
fibrinogen alpha chain	gi|458553	e-162	OS
fibrinogen alpha subunit and fibrinogen alpha subunit precursor	gi|182406	7e-38	OS
fibrinogen alpha subunit and fibrinogen alpha subunit precursor	gi|182406	e-109	OS
fibrinogen alpha subunit and fibrinogen alpha subunit precursor	gi|182406	0	OS
fibrinogen alpha-chain	gi|182423	2e-59	OS
fibrinogen beta-chain	gi|182429	0	OS
fibrinogen beta-chain	gi|182429	e-178	OS
fibrinogen beta-chain	gi|182429	e-175	OS
fibrinogen beta-chain	gi|182429	e-143	OS
fibrinogen beta-chain	gi|182429	0	OS
fibrinogen beta-chain	gi|182429	2e-69	OS
fibrinogen beta-chain	gi|182429	4e-79	OS
fibrinogen, A alpha polypeptide (FGA)	gi|4503688	e-101	OS
fibrinogen, gamma	gi|182441	8e-84	OS
fibrinogen, gamma polypeptide (FGG)	gi|4503714	6e-78	OS
fibrinogen, gamma polypeptide (FGG)	gi|4503714	e-159	OS
fibrinogen, gamma polypeptide (FGG)	gi|4503714	0	OS
fibrinogen, gamma polypeptide (FGG)	gi|4503714	0	OS
fibrinogen, gamma polypeptide (FGG)	gi|4503714	e-171	OS
fibrinogen, gamma polypeptide (FGG) mRNA	gi|4503714	e-164	OS

Subject	Accession#	P Value	Class
group-specific component (GC)	gi|31675	e-162	OS
haptoglobin alpha 1S (Hpa 1S)	gi|32429	e-156	OS
haptoglobin alpha 1S (Hpa 1S)	gi|32429	0	OS
haptoglobin alpha 1S (Hpa 1S)	gi|32429	e-162	OS
haptoglobin alpha 1S (Hpa 1S)	gi|32429	e-116	OS
haptoglobin alpha 1S (Hpa 1S)	gi|32429	0	OS
haptoglobin alpha 1S (Hpa 1S)	gi|32429	0	OS
haptoglobin alpha 1S (Hpa 1S)	gi|32429	0	OS
haptoglobin alpha- and beta-chain	gi|32417	1e-06	OS
haptoglobin alpha- and beta-chain	gi|32417	e-172	OS
haptoglobin alpha(1S)-beta precursor	gi|184316	e-144	OS
haptoglobin alpha(1S)-beta precursor	gi|184316	e-140	OS
haptoglobin alpha(1S)-beta precursor	gi|184316	0	OS
haptoglobin alpha(2FS)-beta	gi|459812	e-104	OS
haptoglobin alpha(2FS)-beta	gi|459812	e-153	OS
haptoglobin alpha(2FS)-beta	gi|459812	e-147	OS
haptoglobin alpha(2FS)-beta	gi|459812	0	OS
haptoglobin alpha(2FS)-beta	gi|459812	e-164	OS
haptoglobin alpha(2FS)-beta	gi|459812	e-102	OS
haptoglobin alpha(2FS)-beta	gi|459812	e-129	OS
haptoglobin alpha(2FS)-beta precursor	gi|31749	e-140	OS
haptoglobin alpha(2FS)-beta precursor	gi|31749	e-115	OS
haptoglobin alpha(2FS)-beta precursor	gi|31749	e-171	OS
haptoglobin alpha(2FS)-beta precursor	gi|31749	0	OS
haptoglobin alpha(2FS)-beta precursor	gi|31749	0	OS
haptoglobin alpha(2FS)-beta precursor	gi|31749	0	OS
haptoglobin and haptoglobin-related protein (HP and HPR)	gi|292156	e-140	OS
haptoglobin gene (alpha-2 allele)	gi|184327	0	OS
hemopexin	gi|184487	e-129	OS
histidine-rich glycoprotein	gi|184391	0	OS
inter-alpha-trypsin inhibitor family heavy chain-related protein (IHRP)	gi|664887	e-122	OS
inter-alpha-trypsin inhibitor subunit 3	gi|32605	e-111	OS
lipopolysaccharide binding protein (LBP)	gi|2653815	e-104	OS
microglobulin, alpha-1- (protein HC)	gi|32046	e-169	OS
microglobulin, alpha-1- and HI-30	gi|24478	e-178	OS
plasminogen	gi|190112	e-179	OS
retinol binding protein (RBP)	gi|35896	e-111	OS
retinol binding protein (RBP)	gi|35896	e-117	OS
serum albumin	gi|28591	e-138	OS
serum albumin (HSA)	gi|28589	0	OS
serum albumin (HSA)	gi|28589	0	OS
serum albumin (HSA)	gi|28589	e-162	OS
serum albumin (HSA)	gi|28589	7e-87	OS
serum amyloid A (SAA)	gi|337747	e-166	OS
serum amyloid A (SAA)	gi|1160968	e-119	OS
serum amyloid A (SAA)	gi|1160968	e-112	OS
serum amyloid A (SAA) mRNA.	gi|337747	e-158	OS
serum amyloid A (SAA) protein partial, clone pAS3-alpha	gi|36313	e-105	OS
serum amyloid A1 (SAA1)	gi|758678	e-120	OS
serum amyloid A1 (SAA1)	gi|758678	e-162	OS
serum amyloid A1 (SAA1)	gi|758678	0	OS
serum amyloid P component (SAP)	gi|36330	e-167	OS
serum vitamin D-binding protein (hDBP)	gi|181481	7e-71	OS
thyroid hormone binding protein (p55)	gi|339646	e-147	OS
transferrin	gi|248647	e-168	OS
transferrin [human, liver, mRNA, 2347 nt]	gi|248647	e-172	OS
vascular cell adhesion molecule 1	gi|340193	1e-97	OS
zinc alpha-2-glycoprotein 1 (AZGP1)	gi|4502336	4e-82	OS

Subject	Accession#	P Value	Class
DNA topoisomerase II (partial)	gi|288564	0	TT
heterogeneous nuclear ribonucleoprotein complex K	gi|241477	0	TT
LPS-Induced TNF-Alpha Factor (LITAF)	gi|1684871	e-156	TT
nucleolin (C23) (h)	gi|191026	7e-98	TT
polyadenylate binding protein II	gi|693936	0	TT
ribosomal protein L3	gi|313658	e-158	TT
transcription factor CA150, putative	gi|2460123	e-160	TT

b) The 180 EST sequences of $P < 0.001$ generated from 200 clones of human adult liver tumor cDNA library.

Subject	Accession#	P Value	Class
80K-L protein	gi|219893	e-140	CS
ADP-ribosylation factor 1 , clone 24537	gi|3360490	e-172	CS
annexin IV (carbohydrate-binding protein p33/41)	gi|1060889	e-101	CS
annexin IV (carbohydrate-binding protein p33/41)	gi|1060889	0	CS
calreticulin (CALR)	gi|179881	4e-54	CS
nm23-H2	gi|35069	e-121	CS
ornithine decarboxylase (ODC)	gi|189372	e-133	CS
small G protein RIBB	gi|2286100	e-171	CS
thymosin beta-4	gi|339688	e-101	CS
tyrosine kinase, non-receptor protein	gi|37503	e-139	CS
dynein light chain 1, cytoplasmic (hdlc1)	gi|1209060	6e-97	CY
keratin 18	gi|186690	1e-66	CY
profilin II	gi|190387	e-111	CY
tubulin, beta-	gi|338694	e-169	CY
cytochrome c oxidase subunit VIb (EC 1.9.3.1)	gi|30294	e-138	EN
cytochrome oxidase subunit VIb	gi|311347	e-160	EN
electron transfer flavoprotein beta subunit	gi|297901	e-150	EN
enolase, non-neuronal	gi|31176	1e-97	EN
NADH:ubiquinone oxidoreductase B15 subunit	gi|4164445	e-152	EN
ubiquinol-cytochrome c reductase, smallest subunit of	gi|2317645	e-150	EN
tissue inhibitor of metalloproteinase-3 precursor (TIMP-3)	gi|472309	e-141	EX
EBI3-associated protein p60	gi|1145798	e-163	H
adipophilin	gi|1806039	e-134	M
murine tumor rejection antigen gp96, tra1 mRNA for human homologue of	gi|37260	2e-92	M
signal sequence receptor, beta-	gi|1019367	e-108	M
tetraspan transmembrane protein L6H, putative (TM4SF5)	gi|2587053	e-159	M
CI-MNLL homolog gene	gi|3342003	e-118	MI
D10S102=FBRNP	gi|399757	e-102	MI
dek	gi|30502	e-175	MI
fragile X mental retardation protein 1 homolog FXR1	gi|887792	e-169	MI
H19 gene	gi|3789930	e-148	MI
heat shock 10kD protein 1 (chaperonin 10) (HSPE1)	gi|469170	e-126	MI
hypothetical protein, clone YR-29	gi|3881975	e-100	MI
immunoglobin light chain, (lambda) DNA for	gi|2114261	5e-23	MI
KIAA0103	gi|285942	e-147	MI
non-histone chromosomal protein (HMG-1) retropseudogene.	gi|184250	3e-80	MI
nucleolar protein p40	gi|1835785	e-159	MI
ubiquitin A-52 residue ribosomal protein fusion product 1 (UBA52)	gi|37564	4e-76	MI
xs99	gi|533965	e-119	MI
cytochrome oxidase subunit II (COII), Hsa6 mitochondrion	gi|530076	0	MITO

Subject	Accession#	P Value	Class
Isolate Aus4 cytochrome b (CYTB) gene, mitochondrial gene encoding mitochondrial protein, partial cds	gi|3372411	e-174	MITO
mitochondrial DNA	gi|644480	1e-78	MITO
mitochondrial DNA, complete genome	gi|1262342	e-168	MITO
mitochondrial DNA, complete genome	gi|1262342	e-169	MITO
mitochondrial DNA, complete genome	gi|1262342	e-132	MITO
mitochondrial DNA, complete genome	gi|1262342	e-135	MITO
mitochondrial DNA, complete genome	gi|1262342	9e-99	MITO
mitochondrial DNA, complete genome	gi|1262342	0	MITO
mitochondrial DNA, complete genome	gi|1262342	e-153	MITO
mitochondrial DNA, complete sequence	gi|644480	2e-61	MITO
mitochondrial genes for several tRNAs (Phe, Val, Leu) and 12S and 16S ribosomal RNAs	gi|13683	5e-60	MITO
mitochondrial genes for tRNA (Phe) and 12S rRNA (fragment)	gi|12998	e-112	MITO
mitochondrial genes for tRNA (Phe) and 12S rRNA (fragment)	gi|12998	e-104	MITO
mitochondrial genome	gi|13003	2e-97	MITO
mitochondrial genome	gi|13003	e-155	MITO
mitochondrial genome (consensus sequence)	gi|13000	5e-78	MITO
mitochondrial genome (consensus sequence)	gi|13000	e-148	MITO
mitochondrion, complete genome	gi|1944628	e-158	MITO
NADH dehydrogenase subunit 2 (ND2) gene, mitochondrial gene encoding mitochondrial protein	gi|3287315	e-165	MITO
novel	gi|458031	5e-05	N
Novel	gi|1272369	1e-06	N
novel (12q24.2 PAC RPCI1-128M12)	gi|3366553	4e-71	N
novel (BAC clone GS180J05 from 7q11)	gi|2388557	2e-37	N
novel (BAC clone RG313E03 from 7q31)	gi|2828785	e-159	N
novel (BTK region clone 2f10-rpi)	gi|460086	6e-85	N
novel (chromosome 21 derived BAC containing erythroblastosis virus oncogene homolog 2 protein (ets-2) gene)	gi|2736086	e-164	N
novel (Chromosome 22q11.2 PAC Clone p_n5 In BCRL2-GGT Region)	gi|3900814	9e-23	N
novel (chromosome 5, BAC clone 259m9 (LBNL H193))	gi|3540154	e-138	N
novel (clone 23884)	gi|3005703	9e-80	N
novel (clone 24560 unknown)	gi|3005718	e-154	N
novel (clone 24896)	gi|4406688	e-128	N
novel (clone 24955)	gi|4406629	e-136	N
novel (clone 557 OTK27)	gi|3859989	e-177	N
novel (clone lambda MEN1 region unknown protein)	gi|2529720	1e-89	N
novel (DNA from cosmid MMDA from chromosome 19q13.3)	gi|188616	7e-78	N
novel (DNA sequence from BAC 445C9 on chromosome 22q12.1)	gi|2661905	5e-66	N
novel (DNA sequence from clone 88L2 on chromosome 1)	gi|4140364	3e-24	N
novel (DNA sequence from clone E146D10 on chromosome 22)	gi|3402828	e-103	N
novel (DNA sequence from PAC 148E22 on chromosome 20q12-q13.12)	gi|2791551	e-103	N
novel (DNA sequence from PAC 368A4 on chromosome X)	gi|2578095	e-150	N
novel (DNA sequence from PAC 368A4 on chromosome X)	gi|2578095	e-105	N
novel (Genomic sequence from Human 9q34)	gi|2133911	e-165	N
novel (Genomic sequence from Human 9q34)	gi|2133894	3e-64	N
novel (PAC clone DJ0777O23 from 7p14-p15)	gi|3242763	4e-62	N
novel (PAC clone DJ0808A01 from 7q21.1-q31.1)	gi|3694662	e-134	N
26S proteasome ATPase subunit	gi|2791679	e-176	OM
aldehyde dehydrogenase 1 (ALDH1)	gi|2183298	e-168	OM
aldolase B	gi|178352	4e-88	OM

Subject	Accession#	P Value	Class
ferritin H chain	gi|28434	e-150	OM
ferritin L chain	gi|2230868	2e-62	OM
ferritin L chain	gi|2230868	e-131	OM
ferritin L chain	gi|182513	e-141	OM
ferritin L chain	gi|2230868	e-135	OM
ferritin L chain	gi|182513	e-161	OM
ferritin L chain	gi|28518	e-134	OM
monoamine oxidase A (MAOA)	gi|187352	3e-80	OM
monocyte/macrophage serine-esterase 1, partial	gi|36421	0	OM
mutant cerebroside sulfate activator protein (SAP-MU-0) mRNA, complete cds, without the 9 bp insertion found inalternatively spliced mRNA.	gi|337764	e-124	OM
Na+/H+ exchange regulatory co-factor (NHERF)	gi|2920584	e-117	OM
NADPH-flavin reductase	gi|1384067	e-168	OM
natural killer cell enhancing factor (NKEFB)	gi|440307	e-137	OM
NDUFV3 gene	gi|2230875	6e-63	OM
proteasome subunit HsN3	gi|565650	2e-97	OM
proteasome subunit z	gi|1531532	e-128	OM
sorbitol dehydrogenase (SORD)	gi|2228516	e-152	OM
ubiquitin	gi|1490418	e-170	OM
antitrypsin, alpha 1-	gi|28965	4e-95	OS
antitrypsin, alpha-1-	gi|177828	e-154	OS
apolipoprotein AII	gi|28743	e-142	OS
apolipoprotein A-II	gi|178423	e-138	OS
apolipoprotein A-II (APOA2)	gi|4502148	e-161	OS
apolipoprotein A-II (APOA2)	gi|4502148	e-102	OS
apolipoprotein A-II (APOA2)	gi|4502148	e-155	OS
apolipoprotein A-II (APOA2)	gi|4502148	e-122	OS
apolipoprotein A-II (APOA2)	gi|4502148	2e-37	OS
apolipoprotein A-II (APOA2)	gi|4502148	e-168	OS
apolipoprotein A-II (APOA2)	gi|4502148	e-137	OS
apolipoprotein A-II (APOA2)	gi|4502148	0	OS
apolipoprotein A-II (APOA2)	gi|4502148	1e-67	OS
apolipoprotein A-II (APOA2)	gi|4502148	0	OS
apolipoprotein A-II (APOA2)	gi|4502148	e-141	OS
apolipoprotein CII	gi|28799	e-113	OS
apolipoprotein CII	gi|28799	2e-90	OS
apolipoprotein CII	gi|28805	e-169	OS
apolipoprotein C-II (APOC2)	gi|4502158	5e-63	OS
apolipoprotein C-III (apoC-III)	gi|178828	e-177	OS
haptoglobin Hp-alpha-2	gi|184322	e-101	OS
microglobulin, alpha-1- (protein HC)	gi|32046	0	OS
microglobulin, alpha-1- (protein HC)	gi|32046	e-174	OS
plasminogen	gi|190112	0	OS
prealbumin	gi|189583	e-137	OS
prealbumin	gi|189581	0	OS
retinol binding protein (RBP)	gi|35896	7e-38	OS
serum albumin	gi|28591	e-152	OS
serum albumin (ALB)	gi|4502026	e-152	OS
serum albumin (ALB)	gi|178343	2e-69	OS
serum albumin (ALB)	gi|4502026	0	OS
thyroid hormone binding protein (p55)	gi|339646	0	OS
acidic ribosomal phosphoprotein P0	gi|190231	e-126	TT
acidic ribosomal phosphoprotein P1	gi|190233	e-156	TT
DNA-binding protein, insulinoma rig-analog mRNA encoding	gi|184553	1e-82	TT
H3.3A variant histone (b)	gi|1022324	e-144	TT
nuclear factor NF45	gi|532312	2e-93	TT
ribosomal protein L10	gi|414586	e-129	TT
ribosomal protein L10	gi|414586	e-150	TT

Subject	Accession#	P Value	Class
ribosomal protein L18 (RPL18)	gi|337492	e-128	TT
Ribosomal protein L18a	gi|401844	e-152	TT
ribosomal protein L18a	gi|401844	0	TT
ribosomal protein L23	gi|34198	e-110	TT
ribosomal protein L23a	gi|1399085	e-129	TT
ribosomal protein L27 (RPL27)	gi|388768	e-175	TT
ribosomal protein L27a	gi|550016	e-110	TT
ribosomal protein L28	gi|550018	3e-70	TT
ribosomal protein L32	gi|36131	3e-98	TT
ribosomal protein L35	gi|562073	3e-73	TT
ribosomal protein L37	gi|1839333	7e-84	TT
ribosomal protein L37a	gi|292438	4e-79	TT
ribosomal protein L7	gi|36139	e-107	TT
ribosomal protein L8	gi|433898	0	TT
ribosomal protein L8	gi|433898	e-128	TT
ribosomal protein S16	gi|338446	e-116	TT
ribosomal protein S18	gi|38422	e-104	TT
Ribosomal protein S19	gi|337732	2e-53	TT
ribosomal protein S19	gi|337732	e-172	TT
ribosomal protein S19	gi|337732	e-130	TT
ribosomal protein S20 (RPS20)	gi|292442	e-118	TT
ribosomal protein S20 (RPS20)	gi|292442	e-134	TT
ribosomal protein S20 (RPS20)	gi|292442	e-121	TT
ribosomal protein S20 (RPS20)	gi|292442	0	TT
ribosomal protein S20 (RPS20)	gi|292442	e-103	TT
ribosomal protein S21 (RPS21)	gi|292444	e-106	TT
Ribosomal protein S21 (RPS21)	gi|292444	e-138	TT
ribosomal protein S27	gi|431318	2e-96	TT
ribosomal protein S27	gi|1373420	e-167	TT
ribosomal protein S28	gi|388030	e-166	TT
ribosomal protein s3, 40S	gi|32531	e-144	TT
ribosomal RNA, 16S	gi|2245558	e-113	TT
translation elongation factor 1-alpha (EF1A)	gi|31091	e-128	TT
translation elongation factor-1-beta	gi|31099	e-160	TT
TRAP-complex gamma subunit (a)	gi|312701	5e-21	TT

Key:

(a) Rattus norvegicus (b) Mus musculus (c) Oryctolagus cuniculus (d) Arabidopsis thaliana (e) Caenorhabditis elegans (f) Nicotiana tabacum (g) Drosophila melangaster (h) Chinese hamster *(I) Oryza sativa (j) Ipomoea batatas (k) Anabaena sp.*

3.2 Nomenclature of human NLT gene

The investigation of organic transporters has been started since the last decade. To date, there are at least two multispecific organic transporter families that have been identified, while the discovery of organic anion transporters are the latest finding in these few years. In order to prevent confusion or mix up with other anion transporters such as organic anion transporter polypeptide (oatp) or canalicular multispecific organic anion transporter (cMOAT), a formal and systemic name is necessary for the gene.

The name NLT that stands for novel liver-specific transporter was first assigned when a novel rat liver-specific transporter was being identified (Simonson et al., 1994). In additional to the NLT amino acid sequence that exhibits 42% identity to that of OAT1, it was also further demonstrated that it has multispecific organic anion transport properties by Sekine *et al.* (1998) and a name called OAT2 was being proposed. Until 1999, Kim *et al.* proposed the corresponding provisional human sequence, which showed 78% amino acid identity. In order to have a centralized numbering system to distinguish the individual members of a human gene family, a new gene symbol was necessarily designated, in consideration of its structure, function and expression data. Having approved and suggested by the HUGO Gene Nomenclature Committee (HGNC),

a formal gene name was denoted by our group to NLT or OAT2 as the Solute Carrier Family 22 member 7 (SLC22A7).

Up till now, the name OAT2 is commonly described in published papers. Together with the other three members hOAT1, hOAT3 and hOAT4, they are all belong to a new subset of multispecific organic anion transporter family other than ABC and OATP family, or known as amphiphilic solute facilitator (ASF) family (Race *et al.*, 1999).

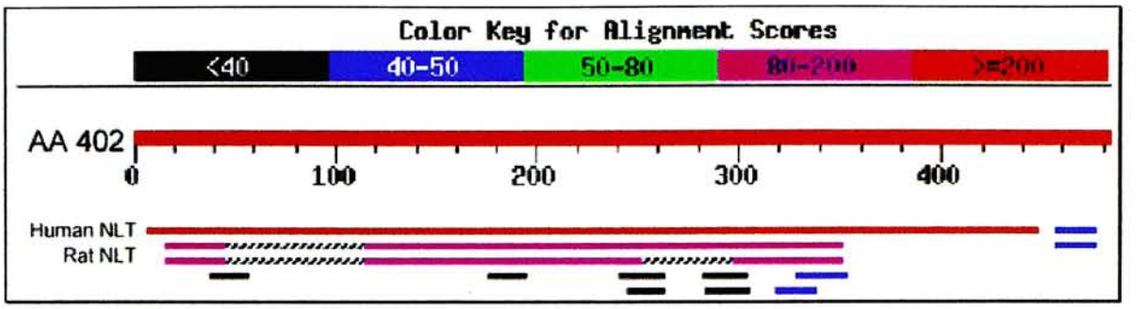
3.3 Cloning and characterization of the hOAT2 sequence

3.3.1 Isolation of hOAT2 cDNA from human liver cDNA library

From the large pool of human normal liver EST sequences generated previously in our laboratory, a partial cDNA clone named aa402 (483 bp) as one of the novel genes was identified. The sequencing pattern is shown in figure 3.2. Under the non-redundant BLAST search from NCBI, two reported sequences were shown apparently to match closely to the clone. They are the human liver-specific transporter (NLT) (Accession: AF097518) from Kim *et al.* (1999, unpublished), and the rat liver-specific transporter (Accession: L27651) from Simonson *et al.* (1994) with nucleotide identities up to 95 % and 88% respectively (Figure 3.1). Although the full length cDNA sequence of the human

liver-specific transporter has already been found, further sequence properties and protein functions associated with the gene have not yet been characterized.

Figure 3.1 The nucleotide sequence alignment of the clone aa402 was found to match with high identity with human and rat NLT under the BLAST search.




```
                                                                  Score     E
Sequences producing significant alignments:                       (bits)  Value

gb|AF097518.1|AF097518  Homo sapiens liver-specific transpor...   662    0.0
gb|L27651.1|RATLSTP  Rattus norvegicus liver-specific transp...   159    1e-36
gb|L30107.1|RATLIVTR  Rattus norvegicus liver-specific trans...   101    2e-19
gb|AF013754.1|AF013754  Gallus gallus transcription factor R...    42    0.20
gb|L48477.1|HUM21DC93Z  Homo sapiens (subclone 6_b10 from P1...    40    0.78
emb|AL031231.1|SC3C3  Streptomyces coelicolor cosmid 3C3           40    0.78
emb|AL049829.4|CNS0000B  Human chromosome 14 DNA sequence **...    38    3.1
emb|AL021707.2|HS508I15  Human DNA sequence from clone RP3-5...    38    3.1
gb|U05958.1|SANKCC1  Squalus acanthias bumetanide-sensitive ...    38    3.1
emb|X13768.1|SSP45SC  Porcine mRNA for cytochrome P-450 (SCC...    38    3.1
emb|Z96601.1|HS3QT013  H.sapiens telomeric DNA sequence, clo...    38    3.1
emb|X75956.1|BTTGNPFMR  B.taurus mRNA for trans golgi networ...    38    3.1
Alignments

>gb|AF097518.1|AF097518 Homo sapiens liver-specific transporter (NLT) mRNA,
complete cds
Length = 2136

Score =  662 bits (334), Expect = 0.0
Identities = 436/457 (95%), Gaps = 16/457 (3%)
Strand = Plus / Plus

Query: 8   ggcacgctcagctcctgcctccgctttgcctatccccaggcgctctccaacaccacgttg 67
           ||||||||||||||||||||||||||||||||||||||||| ||| ||||||||||||
Sbjct: 318 ggcacgctcagctcctgcctccgctttgcctatccccaggctctccccaacaccacgttg 377

Query: 68  ggggaagaaaggcagagccgtgtgtagctggaggatgaacctgccacagtgccctgctct 127
           |||||||||||||||||||| | ||||||||||||||||||||||||||||||||||
Sbjct: 378 ggggaagaaaggcagagccgtggggagctggaggatgaacctgccacagtgccctgctct 437

Query: 128 cagggctgggagtacgaccactcagaattcttctctaccattgcaactg------agtgg 181
           |||||||||||||||||||||||||||||| ||||||||||||||||      |||||
Sbjct: 438 cagggctgggagtacgaccactcagaattctcctctaccattgcaactgagtcccagtgg 497


Query: 182 gatctggtgtgtgagcagaaaggtctgaacagagctgcgtccactttcttcttcgccggt 241
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 498 gatctggtgtgtgagcagaaaggtctgaacagagctgcgtccactttcttcttcgccggt 557
```

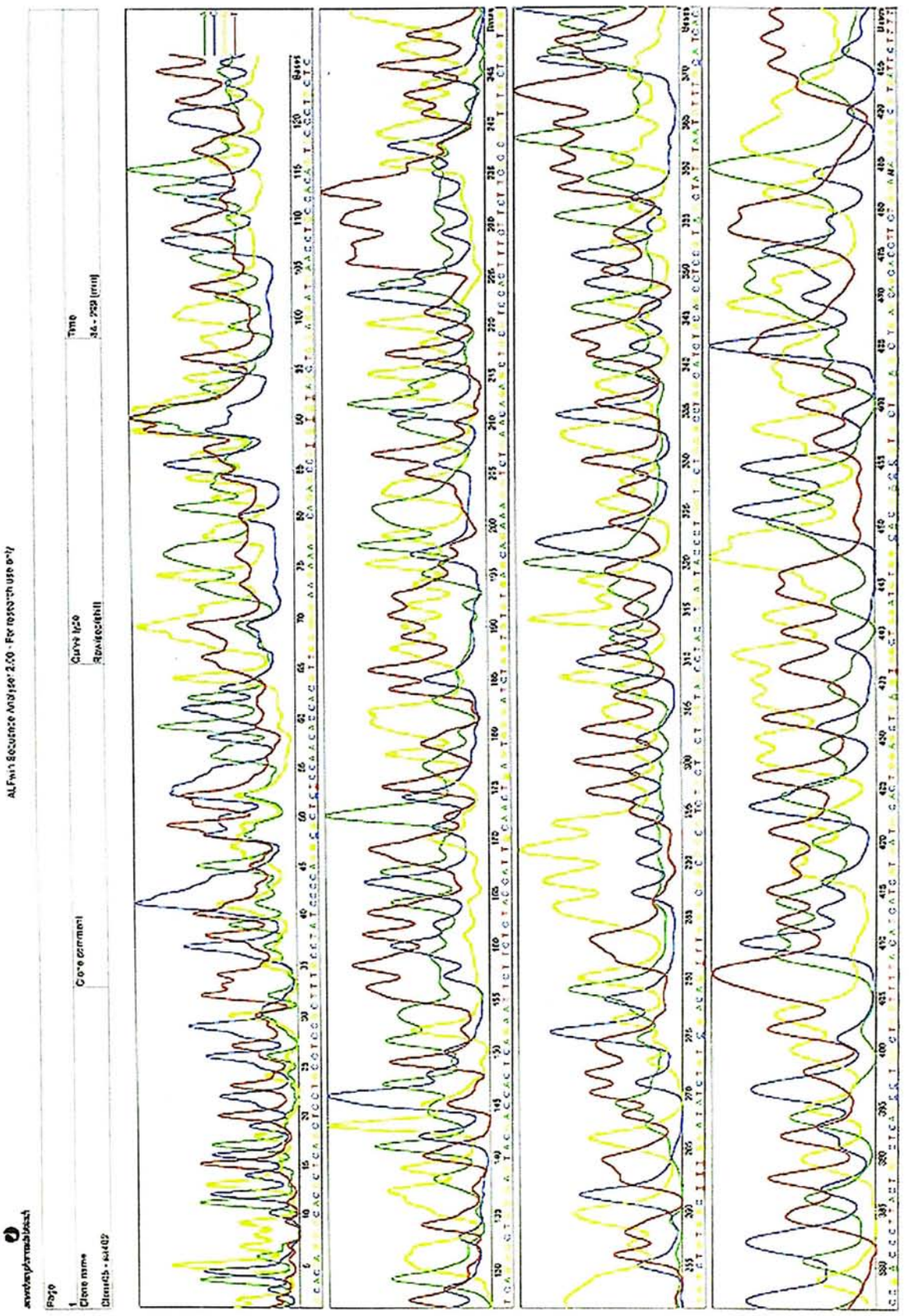
```
Query: 242 gtgctggtgggggctgtgg-ctttggatatctgt-cgaca-gtttgggcggcggcgtctg 298
           |||||||||||||||||| |||||||||||||| ||||| ||||||||||||||||||
Sbjct: 558 gtgctggtgggggctgtggcctttggatatctgtccgacaggtttgggcggcggcgtctg 617

Query: 299 ctgctggtagcctacgtgagtaccctggtgctgggcctggcatctgcagcctccgt-agc 357
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct: 618 ctgctggtagcctacgtgagtaccctggtgctgggcctggcatctgcagcctccgtcagc 677

Query: 358 tatgtaatgtttg-catcacccg-acccttactggctcag-cctggctggtttta-catc 413
           |||||||||||| ||||||||| |||||||||||||||| |||||||||||| ||||
Sbjct: 678 tatgtaatgtttgccatcacccgcacccttactggctcagccctggctggttttaccatc 737

Query: 414 atcgtgatg-cactggagct-gagtggctggatgtgg 448
           ||||||||| |||||||||| ||||||||||||||||
Sbjct: 738 atcgtgatgccactggagctggagtggctggatgtgg 774
```

Figure 3.2 The sequencing result of the partial clone aa402 that was performed by ALF automated sequencer. 483 upright bases were read from the sequence.



3.3.2 The primary and secondary structural analysis of hOAT2

The 1647-nucleotide long hOAT2 cDNA contains an open reading frame that started from 96 and ended at 1742 from the mRNA sequence, and encodes a 549-amino acid residue protein (Figure 3.3). The protein was predicted with a molecular mass of 60 kDa, and it appears to be slightly acidic with isoelectric point (pI) 6.60 as determined by a computer software DNASIS.

An analysis of the primary amino acid sequence of hOAT2 using the Kyte and Doolittle (1982) algorithm predicts twelve α-helical transmembrane domains (Figure 3.4). Similar transmembrane domains pattern was obtained when using the Hoop and Woods model. There is a long extracellular loop consisting of 101 amino acids between the transmembrane domain 1 and 2. The nonpolar rich residues with bulky side chains (Figure 3.5) such as leucine (17%), alanine (10%), valine (8%) and glycine (8%) made the protein highly hydrophobic, which are responsible in promoting the hydrophobic interaction within protein structures. The protein secondary structure was also predicted by Chou and Fasman method provided in the program DNASIS (Figure 3.6), together with the hydropathy plot, they are useful tools in designing immunogenic peptide for raising antibodies and it will be described in the later chapter.

Figure 3.3 Nucleotide and amino acid sequences of the hOAT2 cDNA from the EMBL/GenBank accession number AF097518. The start and stop codons are shown in bold and asterisked respectively.

```
           9         18         27         36         45         54
5'  ATG GGC TTT GAG GAG CTG CTG GAG CAG GTG GGC GGC TTT GGG CCC TTC CAA CTG
    Met Gly Phe Glu Glu Leu Leu Glu Gln Val Gly Gly Phe Gly Pro Phe Gln Leu

          63         72         81         90         99        108
    CGG AAT GTG GCA CTG CTG GCC CTG CCC CGA GTG CTG CTA CCA CTG CAC TTC CTC
    Arg Asn Val Ala Leu Leu Ala Leu Pro Arg Val Leu Leu Pro Leu His Phe Leu

         117        126        135        144        153        162
    CTG CCC ATC TTC CTG GCT GCC GTG CCT GCC CAC CGA TGT GCC CTG CCG GGT GCC
    Leu Pro Ile Phe Leu Ala Ala Val Pro Ala His Arg Cys Ala Leu Pro Gly Ala

         171        180        189        198        207        216
    CCT GCC AAC TTC AGC CAT CAG GAT GTG TGG CTG GAG GCC CAT CTT CCC CGG GAG
    Pro Ala Asn Phe Ser His Gln Asp Val Trp Leu Glu Ala His Leu Pro Arg Glu

         225        234        243        252        261        270
    CCT GAT GGC ACG CTC AGC TCC TGC CTC CGC TTT GCC TAT CCC CAG GCT CTC CCC
    Pro Asp Gly Thr Leu Ser Ser Cys Leu Arg Phe Ala Tyr Pro Gln Ala Leu Pro

         279        288        297        306        315        324
    AAC ACC ACG TTG GGG GAA GAA AGG CAG AGC CGT GGG GAG CTG GAG GAT GAA CCT
    Asn Thr Thr Leu Gly Glu Glu Arg Gln Ser Arg Gly Glu Leu Glu Asp Glu Pro

         333        342        351        360        369        378
    GCC ACA GTG CCC TGC TCT CAG GGC TGG GAG TAC GAC CAC TCA GAA TTC TCC TCT
    Ala Thr Val Pro Cys Ser Gln Gly Trp Glu Tyr Asp His Ser Glu Phe Ser Ser

         387        396        405        414        423        432
    ACC ATT GCA ACT GAG TCC CAG TGG GAT CTG GTG TGT GAG CAG AAA GGT CTG AAC
    Thr Ile Ala Thr Glu Ser Gln Trp Asp Leu Val Cys Glu Gln Lys Gly Leu Asn

         441        450        459        468        477        486
    AGA GCT GCG TCC ACT TTC TTC TTC GCC GGT GTG CTG GTG GGG GCT GTG GCC TTT
    Arg Ala Ala Ser Thr Phe Phe Phe Ala Gly Val Leu Val Gly Ala Val Ala Phe

         495        504        513        522        531        540
    GGA TAT CTG TCC GAC AGG TTT GGG CGG CGG CGT CTG CTG CTG GTA GCC TAC GTG
    Gly Tyr Leu Ser Asp Arg Phe Gly Arg Arg Arg Leu Leu Leu Val Ala Tyr Val

         549        558        567        576        585        594
    AGT ACC CTG GTG CTG GGC CTG GCA TCT GCA GCC TCC GTC AGC TAT GTA ATG TTT
    Ser Thr Leu Val Leu Gly Leu Ala Ser Ala Ala Ser Val Ser Tyr Val Met Phe

         603        612        621        630        639        648
    GCC ATC ACC CGC ACC CTT ACT GGC TCA GCC CTG GCT GGT TTT ACC ATC ATC GTG
    Ala Ile Thr Arg Thr Leu Thr Gly Ser Ala Leu Ala Gly Phe Thr Ile Ile Val

         657        666        675        684        693        702
    ATG CCA CTG GAG CTG GAG TGG CTG GAT GTG GAG CAC CGC ACC GTG GCT GGA GTC
    Met Pro Leu Glu Leu Glu Trp Leu Asp Val Glu His Arg Thr Val Ala Gly Val
```

```
      711        720        729        738        747        756
CTG AGC AGC ACC TTC TGG ACA GGG GGC GTG ATG CTG CTG GCA CTG GTT GGG TAC
Leu Ser Ser Thr Phe Trp Thr Gly Gly Val Met Leu Leu Ala Leu Val Gly Tyr

      765        774        783        792        801        810
CTG ATA CGG GAC TGG CGA TGG CTT CTG CTA GCT GTC ACC CTG CCT TGT GCC CCA
Leu Ile Arg Asp Trp Arg Trp Leu Leu Leu Ala Val Thr Leu Pro Cys Ala Pro

      819        828        837        846        855        864
GGC ATC CTC AGC CTC TGG TGG GTG CCT GAG TCT GCA CGC TGG CTT CTG ACC CAA
Gly Ile Leu Ser Leu Trp Trp Val Pro Glu Ser Ala Arg Trp Leu Leu Thr Gln

      873        882        891        900        909        918
GGC CAT GTG AAA GAG GCC CAC AGG TAC TTG CTC CAC TGT GCC AGG CTC AAT GGG
Gly His Val Lys Glu Ala His Arg Tyr Leu Leu His Cys Ala Arg Leu Asn Gly

      927        936        945        954        963        972
CGG CCA GTG TGT GAG GAC AGC TTC AGC CAG GAG GCT GTG AGC AAA GTG GCC GCC
Arg Pro Val Cys Glu Asp Ser Phe Ser Gln Glu Ala Val Ser Lys Val Ala Ala

      981        990        999       1008       1017       1026
GGG GAA CGG GTG GTC CGA AGA CCT TCA TAC CTA GAC CTG TTC CGC ACA CCA CGG
Gly Glu Arg Val Val Arg Arg Pro Ser Tyr Leu Asp Leu Phe Arg Thr Pro Arg

     1035       1044       1053       1062       1071       1080
CTC CGA CAC ATC TCA CTG TGC TGC GTG GTG GTG TGG TTC GGA GTG AAC TTC TCC
Leu Arg His Ile Ser Leu Cys Cys Val Val Val Trp Phe Gly Val Asn Phe Ser

     1089       1098       1107       1116       1125       1134
TAT TAC GGC CTG AGT CTG GAT GTG TCG GGG CTG GGG CTG AAC GTG TAC CAG ACA
Tyr Tyr Gly Leu Ser Leu Asp Val Ser Gly Leu Gly Leu Asn Val Tyr Gln Thr

     1143       1152       1161       1170       1179       1188
CAG CTG TTG TTC GGG GCT GTG GAA CTG CCC TCC AAG CTG CTG GTC TAC TTG TCG
Gln Leu Leu Phe Gly Ala Val Glu Leu Pro Ser Lys Leu Leu Val Tyr Leu Ser

     1197       1206       1215       1224       1233       1242
GTG CGC TAC GCA GGA CGC CGC CTC ACG CAA GCC GGG ACA CTG CTG GGC ACG GCC
Val Arg Tyr Ala Gly Arg Arg Leu Thr Gln Ala Gly Thr Leu Leu Gly Thr Ala

     1251       1260       1269       1278       1287       1296
CTG GCG TTC GGC ACT AGA CTG CTA GTG TCC TCC GAT ATG AAG TCC TGG AGC ACT
Leu Ala Phe Gly Thr Arg Leu Leu Val Ser Ser Asp Met Lys Ser Trp Ser Thr

     1305       1314       1323       1332       1341       1350
GTC CTG GCA GTG ATG GGG AAA GCT TTT TCT GAA GCT GCC TTC ACC ACT GCT TAC
Val Leu Ala Val Met Gly Lys Ala Phe Ser Glu Ala Ala Phe Thr Thr Ala Tyr

     1359       1368       1377       1386       1395       1404
CTG TTC ACT TCA GAG TTG TAC CCT ACG GTG CTC AGA CAG ACA GGG ATG GGG CTG
Leu Phe Thr Ser Glu Leu Tyr Pro Thr Val Leu Arg Gln Thr Gly Met Gly Leu

     1413       1422       1431       1440       1449       1458
ACT GCA CTG GTG GGC CGG CTG GGG GGC TCT TTG GCC CCA CTG GCG GCC TTG CTA
Thr Ala Leu Val Gly Arg Leu Gly Gly Ser Leu Ala Pro Leu Ala Ala Leu Leu

     1467       1476       1485       1494       1503       1512
GAT GGA GTG TGG CTG TCA CTG CCC AAG CTT ACT TAT GGG GGG ATC GCC CTG CTG
```

```
Asp Gly Val Trp Leu Ser Leu Pro Lys Leu Thr Tyr Gly Gly Ile Ala Leu Leu

      1521       1530       1539       1548       1557       1566
GCT GCC GGC ACC GCC CTC CTG CTG CCA GAG ACG AGG CAG GCA CAG CTG CCA GAG
Ala Ala Gly Thr Ala Leu Leu Leu Pro Glu Thr Arg Gln Ala Gln Leu Pro Glu

      1575       1584       1593       1602       1611       1620
ACC ATC CAG GAC GTG GAG AGA AAG AGT GCC CCA ACC AGT CTT CAG GAG GAA GAG
Thr Ile Gln Asp Val Glu Arg Lys Ser Ala Pro Thr Ser Leu Gln Glu Glu Glu

      1629       1638       1647
ATG CCC ATG AAG CAG GTC CAG AAC TAA 3'
Met Pro Met Lys Gln Val Gln Asn ***
```

Figure 3.4 Hydropathy analysis based on the algorithm of Kyte and Doolittle using a window setting of 8. The putative transmembrane spanning domains are numbered in the corresponding regions.

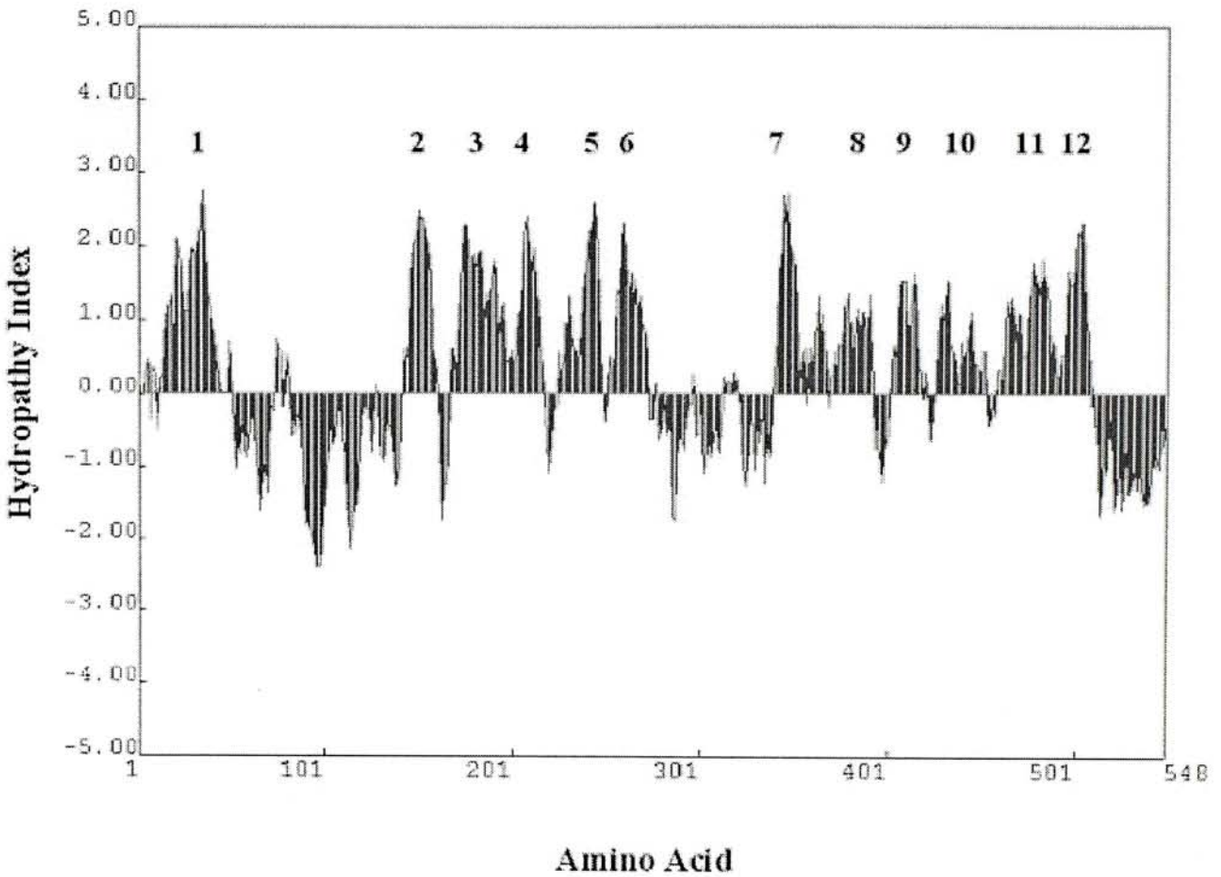


Figure 3.5 The amino acid content of hOAT2. Leucine, alanine and valine accommodate a high percentage of residues among the total 548 amino acids and thus contribute to the hydrophobic property of the protein.

Amino Acid		Count	Mol%
Ala	A	55	10.03
Arg	R	34	6.20
Asn	N	8	1.45
Asp	D	14	2.55
Cys	C	9	1.64
Gln	Q	20	3.64
Glu	E	31	5.65
Gly	G	44	8.02
His	H	10	1.82
Ile	I	10	1.82
Leu	L	92	16.78
Lys	K	9	1.64
Met	M	9	1.64
Phe	F	25	4.56
Pro	P	28	5.10
Ser	S	40	7.29
Thr	T	35	6.38
Trp	W	13	2.37
Tyr	Y	16	2.91
Val	V	46	8.39
Asx	B	0	0.00
Glx	Z	0	0.00
***	*	0	0.00
Xxx	X	0	0.00
Total		548	99.88

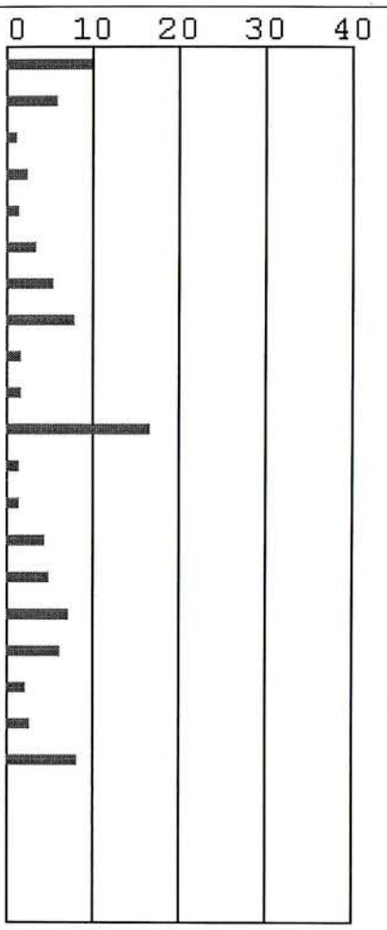


Figure 3.6 The protein secondary structure prediction of hOAT2 by DNASIS using Chou and Fasman method (beta-turn probability value 75 x 10e-1). H(h): helix-turn-helix; S(s): sheet; T(t): turn; and C(c): coil.

```
                10        20        30        40        50        60        70
        MGFEELLEQVGGFGPFQLRNVALLALPRVLLPLHFLLPIFLAAVPAHRCALPGAPANFSHQDVWLEAHLP
HELIX HHHHHHHHhh  h hhhHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHh    hhhHHhhhhHHHHHH
SHEET   ssssssss  s sssssssssssssssssssssssssssss                SSSSs
TURN          TTTTTTTTT                                    TTTTTTTT
COIL

                80        90        100       110       120       130       140
        REPDGTLSSCLRFAYPQALPNTTLGEERQSRGELEDEPATVPCSQGWEYDHSEFSSTIATESQWDLVCEQ
HELIX Hh    hh    hhhhhhhhhhhhhHHHHHh   hHHHHHHH            hhhhHHHHHHHHHHHHh
SHEET       ss    sSSSSSSSssssSs               sss    sSs    sssssssssssssssss
TURN   TTTTTTTTTT       TTTT      TTTTT       TTTTTT TTTT TTTT           T
COIL                                                  C

                150       160       170       180       190       200       210
        KGLNRAASTFFFAGVLVGAVAFGYLSDRFGRRRLLLVAYVSTLVLGLASAASVSYVMFAITRTLTGSALA
HELIX h     HHHhhhhhhhhhhhhhhhhh        hhhhhhhhhhhhhHHHHHHhhhhhhhhhhhhhhhHHH
SHEET         SSSSSSSSSSSSSSSs     sssSSSSSSSSSSSSs     SSSSSSSSSSss   ss
TURN  TTT                     TTTTTTTTT                              TTTT
COIL     CC

                220       230       240       250       260       270       280
        GFTIIVMPLELEWLDVEHRTVAGVLSSTFWTGGVMLLALVGYLIRDWRWLLLAVTLPCAPGILSLWWVPE
HELIX HHHHHHHHHHHHHHHHHHHHHhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhH     hhhhhhhhh
SHEET sssssssssssssss   SSSSSssssSssssSSSSSSSSSSSSSSSSSSSSSsSS   sSSSSSs
TURN                          TTTT TTTT                              TTTT     TTT
COIL

                290       300       310       320       330       340       350
        SARWLLTQGHVKEAHRYLLHCARLNGRPVCEDSFSQEAVSKVAAGERVVRRPSYLDLFRTPRLRHISLCC
HELIX hHHhhhhhhhHHHHHHhhhhHHHh        hhHHHHHHHHHHHHHH                hhhhh
SHEET    SSSssssss     SSSSs                                Ss   sSSSSs     SSSSS
TURN  T     TTTT            TTTTT TTTTTT                  TTTTT    TTTT
COIL                              C                                   CC

                360       370       380       390       400       410       420
        VVVWFGVNFSYYGLSLDVSGLGLNVYQTQLLFGAVELPSKLLVYLSVRYAGRRLTQAGTLLGTALAFGTR
HELIX hhhhhhhhh                hhhhhhhhhhHHhhhhhhhhhhH        hhhhhhhhhhhhhhh
SHEET SSSSSSSss  ssSSSssssSSSSSSSSSSSSSs     SSSSSSsSssssSSSssssSSSSSSSSSS
TURN         TTTTTT   TTTT               TTTT       TTTT   TTTT
COIL

                430       440       450       460       470       480       490
        LLVSSDMKSWSTVLAVMGKAFSEAAFTTAYLFTSELYPTVLRQTGMGLTALVGRLGGSLAPLAALLDGVW
HELIX hhh    hhhhhhhhhHHHHHHHHHhhhhhhhhhhhhhhhhhhh  hhhhhhhhh   hHHHHHHHHHHH
SHEET SSs      SSSSSSSs        SSSSSSSSSSSSSSSSSss  ssSSSSSSs      sssssssss
TURN    TTTTTTT                                  TTTTTT      TTTTT
COIL
```

```
              500         510         520         530         540
       LSLPKLTYGGIALLAAGTALLLPETRQAQLPETIQDVERKSAPTSLQEEEMPMKQVQN
HELIX  HHHHHHhhhhHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH      hHHHHHHHHHHH
SHEET  sssssssssssssssssssssssssssssssssssssssss          ss    sssssss
TURN         TTTT                                    TTTTT
COIL                                                CC              C
```

3.3.3 Motif search and prediction

Protein sequence analysis of hOAT2 from PROSCAN (Bairoch *et al.*, 1997) predicted seven different patterns (Table 3.1). There are three possible *N*-glycosylation sites (N-{P}-[ST]-{P}) at positions Asn^{57}, Asn^{91} and Asn^{358} where {P} indicates that proline is not applicable. Statistical analysis has revealed that the proline residue between the Asn and Ser/Thr will inhibit the N-glycosylation (Bause, 1983). Two of these potential N-glycosylation sites (Asn^{57} and Asn^{91}) are located in the long extracellular loop between transmembrane regions 1 and 2. The carbohydrate chain is attached to the Asn residue if it is in the context of – Asn-X-Ser / Thr- and accessible to the glycosylation machinery localized in the lumen of the ER (Kornfeld and Kornfeld, 1985; Pfeffer and Rothman, 1987). Another interesting finding includes two leucine zipper patterns are deduced whereas this structure often occurs in numerous eukaryotic regulatory proteins and a few in prokaryotic proteins only. The leucine zippers are located near the carboxyl-terminal of the sequence, while the latter one is found within the sequence 461 – 482bp, which is placed within the predicted eleventh and the predicted twelfth membrane-spanning domains. In regard to the leucine zipper pattern often presents in nuclear DNA-binding proteins, it seems to be contradicted if hOAT2 protein is located within the plasma membrane. However,

the high leucine content (totally 92 leu) in this protein probably contributes to such prediction of these leucine-zipper patterns, in other words, this predicted result was just occurred by chance only.

Table 3.3 Protein sequence analysis based on the prediction from PROSCAN, a total of seven different patterns are proposed.

Name	Pattern	Site	
N-glycosylation	N-{P}-[ST]-{P}	57 – 60	NFSH
		91 – 94	NTTL
		358 – 361	NFSY
cAMP- and cGMP-dependent protein kinase phosphorylation	[RK](2)-x-[ST]	330 - 333	RRPS
		402 - 405	RRLT
Protein kinase C phosphorylation	[ST]-x-[RK]	166 - 168	SDR
		281 - 283	SAR
		340 - 342	TPR
		396 - 398	SVR
Casein kinase II phosphorylation	[ST]-x(2)-[DE]	59 - 62	SHQD
		93 - 96	TLGE
		100 - 103	SRGE
		132 - 135	SQWD
		333 - 336	SYLD
		523 - 526	TIQD
		535 - 538	SLQE
N-myristoylation	G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}	53 - 58	GAPANF
		75 - 80	GTLSSC
		142 - 147	GLNRAA
		154 - 159	GVLVGA
		186 - 191	GLASAA
		206 - 211	GSALAG
		233 - 238	GVLSST
		356 - 361	GVNFSY
		370 - 375	GLGLNV
		408 - 413	GTLLGT
		412 - 417	GTALAF
		465 - 470	GMGLTA
		488 - 493	GVWLSL
Amidation	x-G-[RK]-[RK]	169 - 172	FGRR
		400 - 403	AGRR
Leucine zipper pattern	L-x(6)- L-x(6)- L-x(6)- L-x(6)	366 - 387	LDVSGLGLNVY
			QTQLLFGAVEL
		461 – 482	LRQTGMGLTAL
			VGRLGGSLAPL

3.3.4 Homology alignment

Homology matching analyzed by BLAST 2 sequence from NCBI and DNASIS revealed maximum identities between rat and human OAT2 of 82% (1318/1597) in nucleotide sequences and 77% in amino acid sequences (Figure 3.7).

The amino acid sequences of the four members of human organic anion transporter subfamily, including hOAT1 to hOAT4 were compared. The percentage of identity among the sequences is tabulated as below (Table 3.4). There is only a fair identity of hOAT2 to hOAT1, hOAT3 and hOAT4 (38%, 35%, 37% respectively). Instead, the most homologous pair is observed (90%) between hOAT3 and hOAT4, which are located predominantly in kidney and placenta (Cha *et al.*, 2000; Race *et al.*, 1999) respectively.

Table 3.4 Comparison of amino acid sequence identities among the four members in human OAT family

	OAT2	**OAT3**	**OAT4**
OAT1	38%	46%	50%
OAT2		35%	37%
OAT3			90%

Figure 3.7 The amino acid sequence alignment between rat and human OAT2 with 77% (425/541) sequence identity.

```
HOAT2    1 MGFEELLEQVGGFGPFQLRN-VALLALPRVLLPLHFLLPIFLAAVPAHRC   50
           |||| ||  |||||||||||| | | |||| ||| ||||| | |||||| |
ROAT2    1 MGFEDLLDKVGGFGPFQLRNLV-LMALPRMLLPMHFLLPVFMAAVPAHHC   50

HOAT2   51 ALPGAPANFSHQDVWLEAHLPREPDGTLSSCLRFAYPQALPNTTLGEERQ  100
           |||||||| |||| ||||||||| ||  |||||||||  || ||| |
ROAT2   51 ALPGAPANLSHQDLWLEAHLPRETDGSFSSCLRFAYPQTVPNVTLGTEV-  100

HOAT2  101 SR-GELEDEPATVPCSQGWEYDHSEFSSTIATESQWDLVCEQKGLNRAAS  150
           |  || | || |||||||||| |||||||||  ||||| | |||   |
ROAT2  101 SNSGEPEGEPLTVPCSQGWEYDRSEFSSTIATE--WDLVCQQRGLNKITS  150

HOAT2  151 TFFFAGVLVGAVAFGYLSDRFGRRRLLLVAYVSTLVLGLASAASVSYVMF  200
           | || |||||||  |||||||||||||||||| ||||| ||||  | ||
ROAT2  151 TCFFIGVLVGAVVYGYLSDRFGRRRLLLVAYVSSLVLGLMSAASINYIMF  200

HOAT2  201 AITRTLTGSALAGFTIIVMPLELEWLDVEHRTVAGVLSST-FWTGGVMLL  250
             ||||||||||||||| |||||||||||||||| | | || ||| ||
ROAT2  201 VVTRTLTGSALAGFTIIVLPLELEWLDVEHRTVAGVIS-TVFWSGGVLLL  250

HOAT2  251 ALVGYLIRDWRWLLLAVTLPCAPGILSLWWVPESARWLLTQGHVKEAHRY  300
           |||||||| ||||||| |||| ||| | |||||||||||||| | ||  |
ROAT2  251 ALVGYLIRSWRWLLLAATLPCVPGIISIWWVPESARWLLTQGRVEEAKKY  300

HOAT2  301 LLHCARLNGRPVCEDSFSQEA---VSKVAAGERVVRRPSYLDLFRTPRLR  350
           || || |||||| | | ||||   |  |   ||  |||||||||| ||
ROAT2  301 LLSCAKLNGRPVGEGSLSQEALNNV--VTM-ERALQRPSYLDLFRTSQLR  350

HOAT2  351 HISLCCVVVWFGVNFSYYGLSLDVSGLGLNVYQTQLLFGAVELPSKLLVY  400
           ||||||  |||||||||||| |||||||||||||||||||||||||  ||
ROAT2  351 HISLCCMMVWFGVNFSYYGLTLDVSGLGLNVYQTQLLFGAVELPSKIMVY  400

HOAT2  401 -LSVRYAGRRLTQAGTLLGTALAFGTRLLVSSDMKSWSTVLAVMGKAFSE  450
            | ||  ||||| || ||| || ||| ||||   ||| | | | ||||||
ROAT2  401 FL-VRRLGRRLTEAGMLLGAALTFGTSLLVSLETKSWITALVVVGKAFSE  450

HOAT2  451 AAFTTAYLFTSELYPTVLRQTGMGLTALVGRLGGSLAPLAALLDGVWLSL  500
           |||||||||||||||||||| ||||| |||| ||| ||||||||| |
ROAT2  451 AAFTTAYLFTSELYPTVLRQTGLGLTALMGRLGASLARLAALLDGVWLLL  500

HOAT2  501 PKLTYGGIALLAAGTALLLPETRQAQLPETIQDVERKSAPTSLQEEEMPM  550
           ||  |||||| || ||||||||  |||||||||||||  |  |||
ROAT2  501 PKVAYGGIALVAACTALLLPETKKAQLPETIQDVERKS--T--QEED---  550

HOAT2  551 KQVQN.....................................................  600
             |
ROAT2  551 --V*.....................................................  600
```

Figure 3.8 The amino acid sequence of hOAT2 contains 549 residues. The three *N*-glycosylation sites were predicted and labeled as bolded letter "N". The twelve putative membrane-spanning regions are underlined and bolded. In addition, two leucine zipper patterns L-x(6)-L-x(6)-L-x(6)-L-x(6) which labeled in red are identified and the latter one is located within eleventh and twelve transmembrane domains.

NH_3 -

MGFEELLEQV GGFGPFQLRN **VALLALPRVL LPLHFLLPIF**

LAAVPAHRCA LPGAPA**N**FSH QDVWLEAHLP REPDGTLSSC

LRFAYPQALP **N**TTLGEERQS RGELEDEPAT VPCSQGWEYD

HSEFSSTIAT ESQWDLVCEQ KGLNR**AASTF FFAGVLVGAV**

AFGYLSDRFG RRR**LLLVAYV STLVLGLASA ASVSYVMF**AI

TRT**LTGSALA GFTIIVMPLE LEWL**DVEHR**T VAGVLSSTFW**

TGGVMLLALV GYLIRDWR**WL LLAVTLPCAP GILSLWWV**PE

SARWLLTQGH VKEAHRYLLH CARLNGRPVC EDSFSQEAVS

KVAAGERVVR RPSYLDLFRT PRLRH**ISLCC VVVWFGVNFS**

YYGLS*L*DVSG LG***L***N**VYQTQ*L* LFGAVE*L*PSK LLVYL**SVRYA

GRR**LTQAGTL LGTALAFGTR LLVS**SDMKSW ST**VLAVMGKA**

FSEAAFTTAY LFTSELYPTV ***L***RQT**GMG*L*TA LVGR*L*GGSLA**

P*L*AALLDGVW LS**LPKLTYGG IALLAAGTAL LL**PETRQAQL

PETIQDVERK SAPTSLQEEE MPMKQVQN - HOOC

Table 3.5 The corresponding positions of transmembrane helices

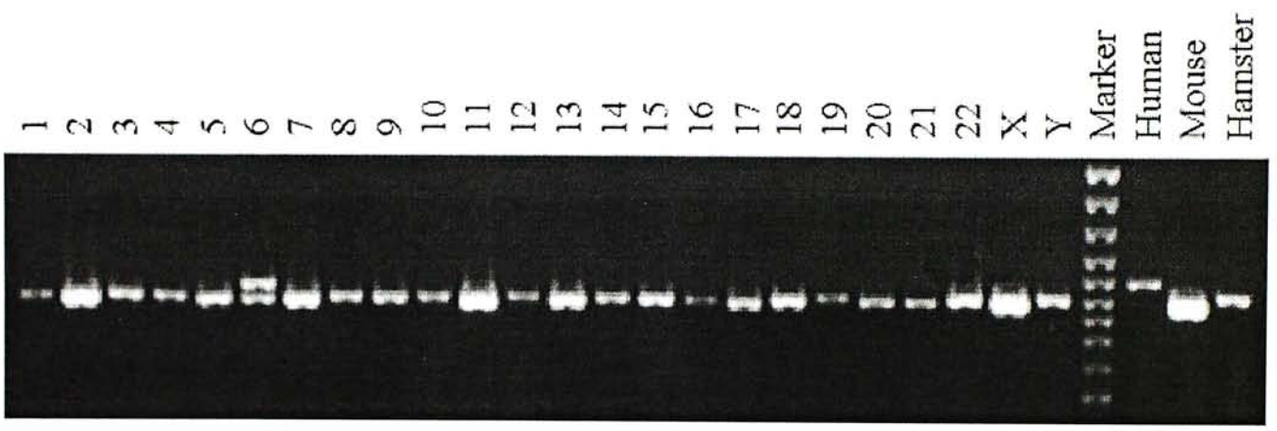
Helix	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Begin	21	146	174	204	230	259	346	375	404	433	465	493
End	44	166	198	224	250	278	366	395	424	453	485	512

3.4 Chromosomal mapping of hOAT2 gene

3.4.1 Somatic cell hybrid mapping of hOAT2 gene

By somatic cell hybrid mapping (monochromosomal NIGMS Human/Rodent Somatic Cell Hybrid Mapping Panel 2-version 3), the chromosomal location of hOAT2 in human was investigated. The PCR products synthesized from somatic cell hybrid contained chromosome (Human line IMR91) 1 to 22, X and Y were loaded on 2% agarose gel. With refer to Figure 3.9, an individual distinct band of about 900 bp, which was amplified from cell hybrid DNA retaining human chromosome 6 (NA10629), was yielded with closest molecular size to human DNA rather than mouse or hamster DNA controls. It indicated the hOAT2 gene was mapped on chromosome 6.

Figure 3.9 PCR screening of a panel of 24 somatic cell hybrid DNAs (human/rodent somatic cell hybrids mapping panel 2-version 3; National Institute of General Medicine Service, Coriell Cell Repositories). The labels indicate on the top of each lane representing the human chromosome number retained in the somatic cell hybrids. Human, mouse and hamster genomic DNA are shown. A band of expected size ~1-kb was generated from chromosome 6 and human genomic DNA control. DNA size marker is DNA Ladder Plus (MBI Fermentas).



3.4.2 Radiation hybrid mapping of hOAT2 gene

For deeper investigation for the location of hOAT2, we performed radiation hybrid mapping containing 93 hybrid genomic DNAs. After analyzing the PCR products by gel electrophoresis (Figure 3.10), we obtained a set of numbers composed of "0" and "1" which represented the non-existence and existence of positive band. The data generated is shown as follow: 10000 00100 00100 00100 01000 10010 11011 11100 01100 10000 00000 11001 01101 00011 00111 01000 11010 00000 000. The screening results were processed by the software RHMAPPER at the Whitehead Institute / MIT Center for Genome Research. The returned output is shown in (Figure 3.11). With reference to the threshold value >15, the hOAT2 gene was found to locate on 0.50cR (LOD>2.5) upper to framework marker D6S271, and 8.01cR lower to framework marker WI-5971, which is equivalent to 249.63 cR from top of Chromosome 6 linkage group. Since the D6S271 marker had already been mapped on 6p21.1 -> p21.2 to which the corresponding location of hOAT2 could be deduced.

Figure 3.10a PCR screening on radiation hybrid panel composed of 93 genomic hybrid DNA from human and hamster somatic cell lines. The samples were analyzed on 2% agarose gel, human and hamster DNA controls were included. The presence of a band which has an identical size to human control is scored as "1", the absence is scored as "0". PCR products amplified from hybrid DNA 1 – 48 were shown.

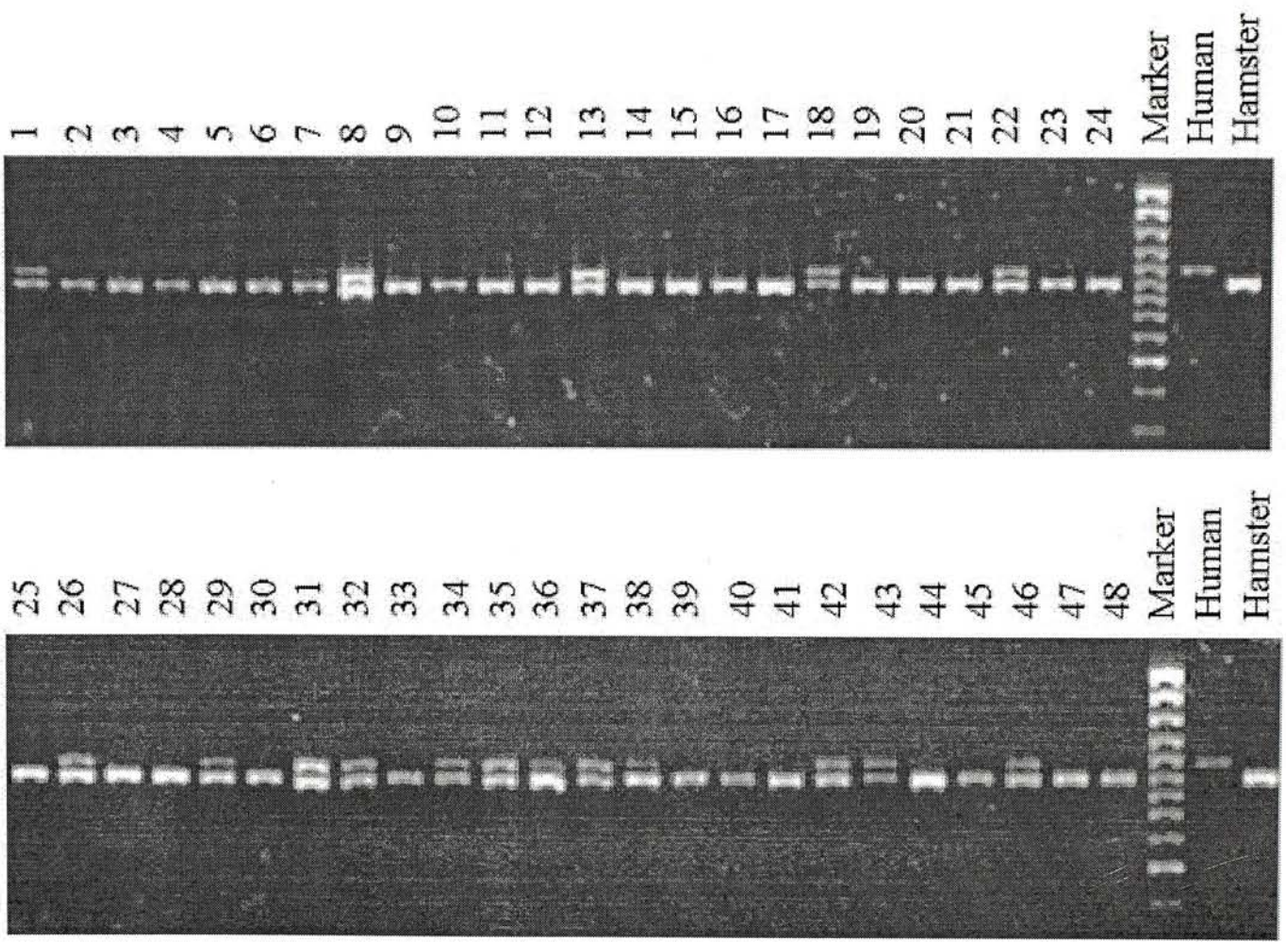


Figure 3.10b In another hybrid DNA panel 49 – 93, sharp and discrete bands were observed.

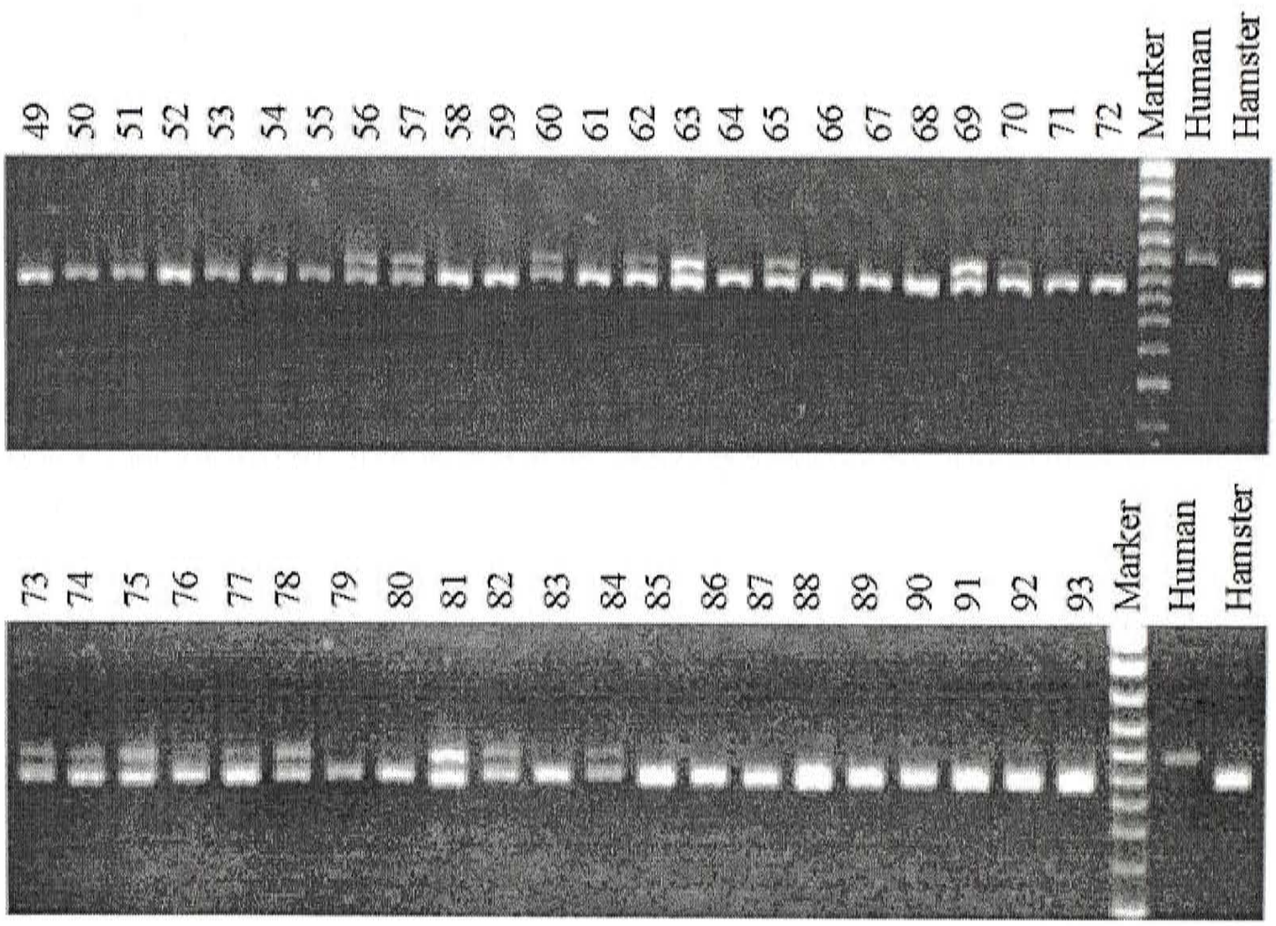
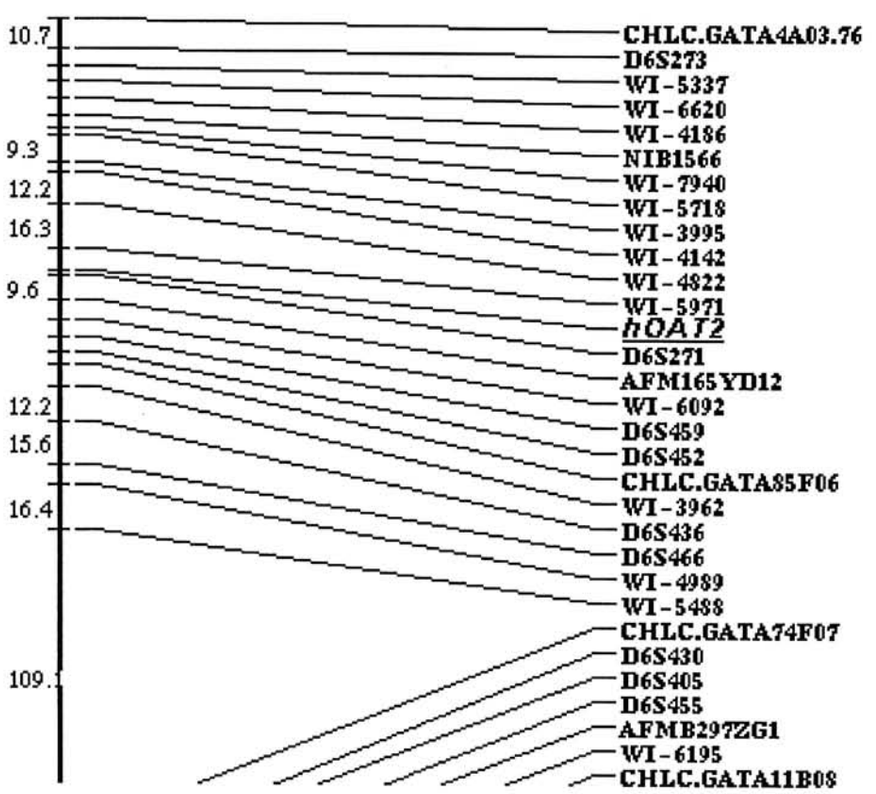


Figure 3.11 A partial RH map from Whitehead Institute showing the placement data relative to framework for human chromosome 6. The hOAT2 was placed within two framework markers, 0.50 cR (LOD > 2.05) upper to D6S271 and 8.01 cR lower to WI-5971. Since the closet position of hOAT2 to the framework marker D6S271, which was mapped 249.63 cR from top of chromosome 6 linkage group, a location of 6p21.1 → 21.2 was identified for hOAT2.



3.4.3 Identification of partial human genomic sequence

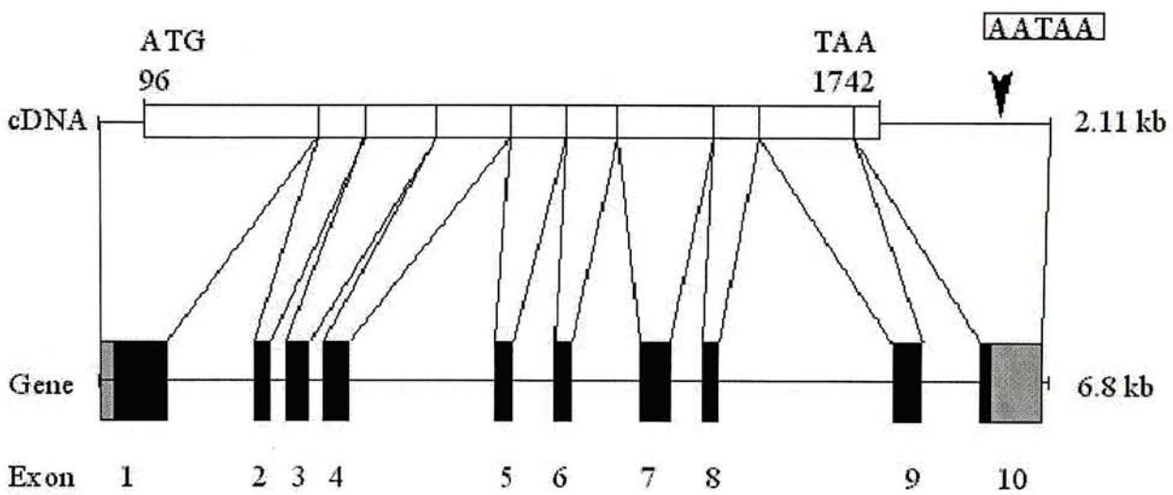
The 1-kb genomic PCR product form the positive human control was subcloned in TA cloning vector and sequenced using M13 reverse primer as described in MATERIALS AND METHODS. It was matched with our expectation that the ~1-kb genomic DNA possessed the cDNA sequence of hOAT2 gene. Concomitantly two introns were identified with size 655-bp and 134-bp respectively and the nucleotide sequences were shown in figure 3.12.

The whole intron-exon distribution of hOAT2 was later deduced after a human genomic clone RP11-117C11 (Accession number: AC026532) was being published. Under the High Throughput Genomic (HTG) Blast search in NCBI, the mRNA sequence of hOAT2 matched in segments of this genomic clone in high identities (percentages). The length of the genomic sequence deduced from the hOAT2 mRNA was approximately 6.8 kb and also nine introns were being deduced accordingly. The former identified two intron sequences were found to be located within exon 2 and 4. The schematic diagram presenting the intron-exon distribution is shown in figure 3.13.

Figure 3.12 The partial genomic sequence of hOAT2 gene. Two introns were identified and sequenced which shows in bold italic letters. The boxes indicated the forward and reverse primers used in the radiation hybrid mapping and the unbold letters represented the coding sequence of hOAT2.

GGCTGGGAGT ACGACCACTC AGAATTCTCC TCTACCATTG CAACTGAG***GT***
ACTTAAGCCC AGAAGTTGAG AGACATGTGA GAGGGATGGG CCTGCCATTT 100
CCTTGGGTGG TTACTGTGTA GCCTATTAGA TGTATTACTT TACCTCTTAA
AGTCTCAGTT TACCAATCTG TAGATGGGGG CTATGAATAT AAGGCACTTT 200
TAGTTCATGG CTTGTAGTAA CAGGATCCTA CACAAGAACT ATTATCATCA
TCATTAATTT GGAGGTGATT AAGGCAGGGA AGGGCTAAAG TGGCTTCAGG 300
GAGTTGAGAC AAAGAGCTGA GAAGGCAATT TCTGCAGCAA GAGGTGGAGA
AACAATAGAG CCCTTCTTTT CTCCCCTTCC TCTTTTCCAA TCCCAGGTCT 400
GTATTTACAT AATCCATCTT GAGGTGGAAT GTCGGTGGAG GCAGTCTCCC
TGGGAGGCAG CAGGTCGAGG CCTTCCTTGG GAAGAAGCTG AGGCTGCAGG 500
ACTGGGGAGG GACAAAGTTT CCTATTCCCC AAGCTGGCGT GAATCGTTGG
GAGGTTTATT ATCTGGCATG GAGGTACCAG AATGGCAGAG TTCGCCTCAG 600
AAGGCTCCAG GGTCTGGAGA GGAGGAGCTG GTGCCTGCTG GAAAGGGGCT
GATGTTCATA GGAGGTCCCT TTCTGCCTGT CCTTGCAGTG GGATCTGGTG 700
TGTGAGCAGA AAGGTCTGAA CAGAGCTGCG TCCACTTTCT TCTTCGCCGG
TGTGCTGGTG GGGGCTGTGG CCTTTGGATA TCTGTCCGAC AGGT***GGGGTG*** 800
AGGCACTGGG CCAATAAGAA ACTGGCTGGG GGAACTTTCT CCCACTAGCT
GGGGTATGAG CCTAGTCTAC CTATGCCTTA GAACCTCCTT CCACAGGGAA 900
CTGACCCTGC ATGACCCCTT TGCAGGTTTG GGCGGCGGCG TCTGCTGCTG
GTAGCCTACG TGAGTACCCT GGTGCTGGGC CT

Figure 3.13 Schematic diagram showing the intron-exon distribution which was determined between the hOAT mRNA sequence and the partial genomic sequence from a *Homo sapiens* clone RP11-117C11 (Accession Number: AC026532). Ten exons were identified and they are shown in black boxes, where the 5' and 3' untranslated regions are indicated in gray.



3.5 Detection of the hOAT2 gene expression in human tissues by RT-PCR assay

A PCR-based analysis was carried out on three human MTC panels: adult, fetal and cancer. PCRs were performed on first-strand cDNA from various human tissues, after that the products were analyzed in 1% agarose gel. For human normal adult panel, sixteen human tissues were chosen in the test. In the figure 3.14, a positive band was highly detected in liver and another band showed less weakly in kidney, no non-specific binding was found in other tissues. This finding was similar to that of rat in which the gene is preferentially expressed in liver and low levels in kidney (Simonson *et al.,* 1994). HOAT2 transcript was not found in heart, brain, placenta, lung, skeletal muscle, pancreas, spleen, thymus, prostate gland, testis, ovary small intestine, colon, or leukocyte even at lower annealing temperature (data not shown).

We have also used another MTC panel in order to determine the expression pattern in human fetus. Eight different human fetal tissues were analyzed in this panel. From this, only a single dominant transcript was detected in fetal liver (Figure 3.15). No positive signal was appeared in fetal kidney, or other fetal tissues including brain, lung, heart, skeletal muscle, spleen or thymus.

The expression of hOAT2 was also found in human liver tumor as shown in

figure 3.16 by amplifying the specific transcript from liver tumor first-strand cDNAs (Invitrogen). Nevertheless, the same pattern was not reported in other carcinomas from human cancer MTC panel (Data not shown).

Figure 3.14 RT-PCR analysis on human normal adult MTC panel demonstrating the tissue distribution for hOAT2. Strong expression was detected in liver relative to kidney to produce a transcript size at ~1.6-kb which was identical to the positive control. Normalization using housekeeping gene GAPDH showed a 550-bp band in all tissues.

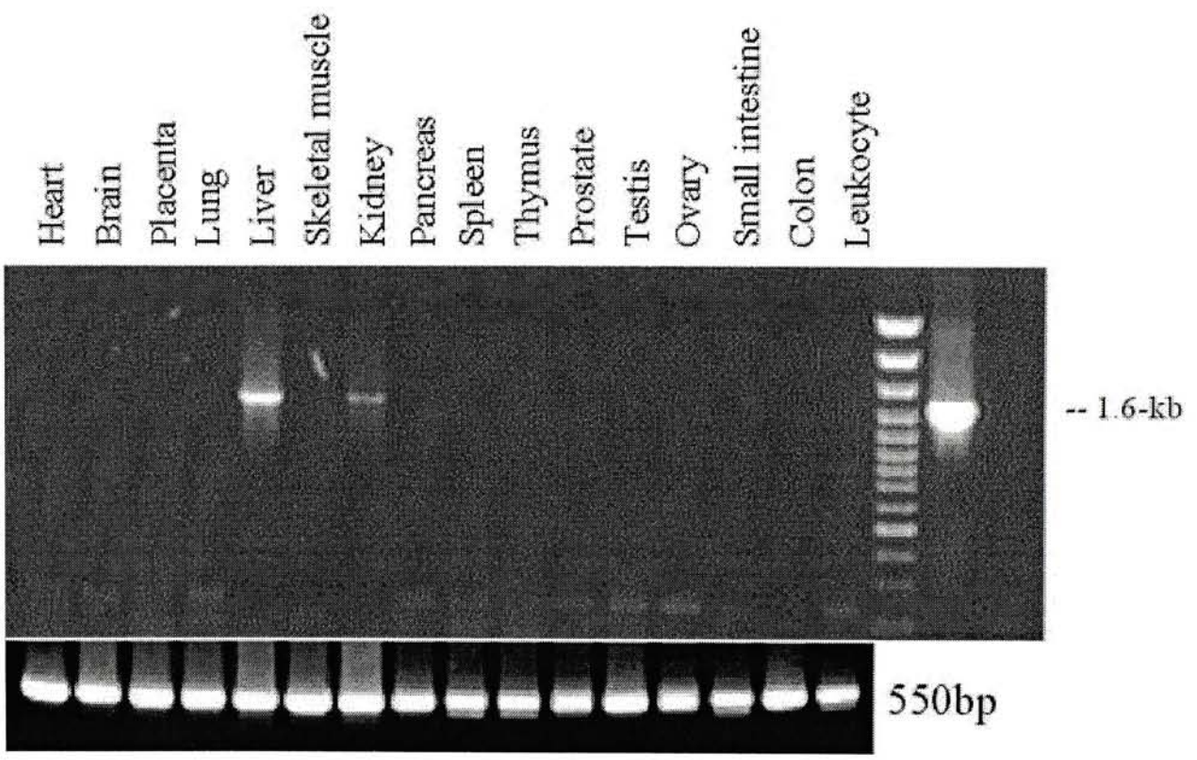


Figure 3.15 RT-PCR analysis on human fetal MTC panel demonstrating the tissue distribution for hOAT2. Expression was only detected in liver producing a transcript size at ~1.3-kb. Normalization was performed using G3PDH which produced a band of ~1-kb in all tissues.

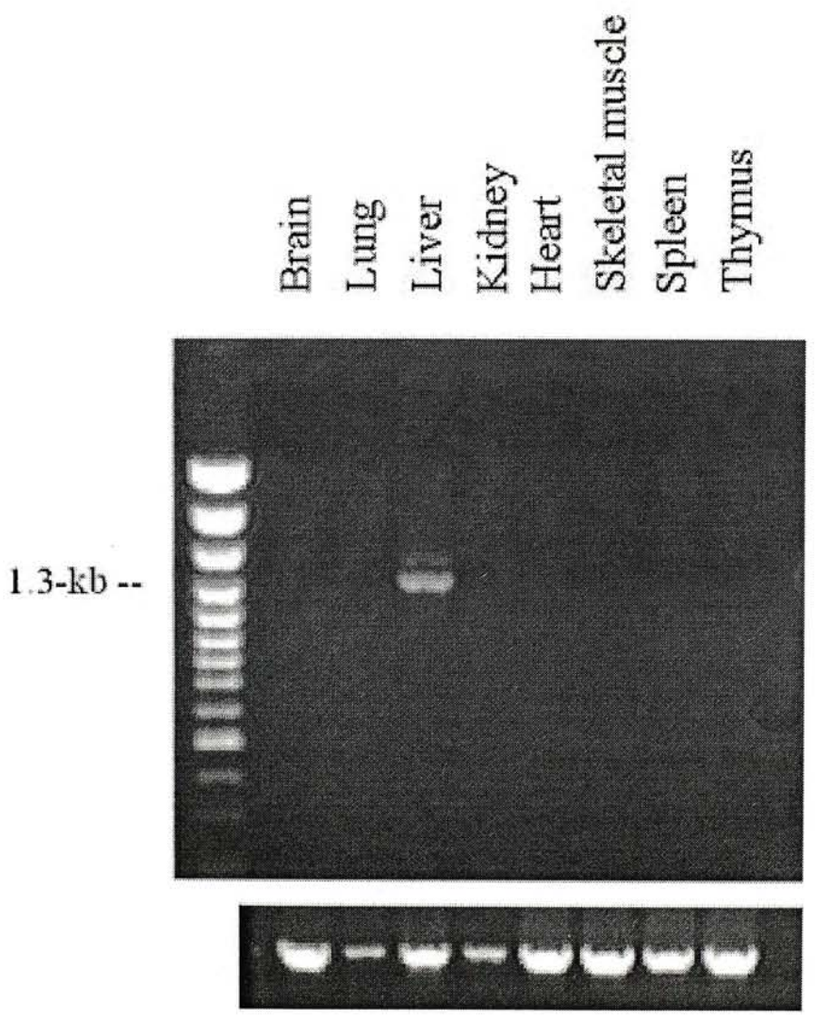
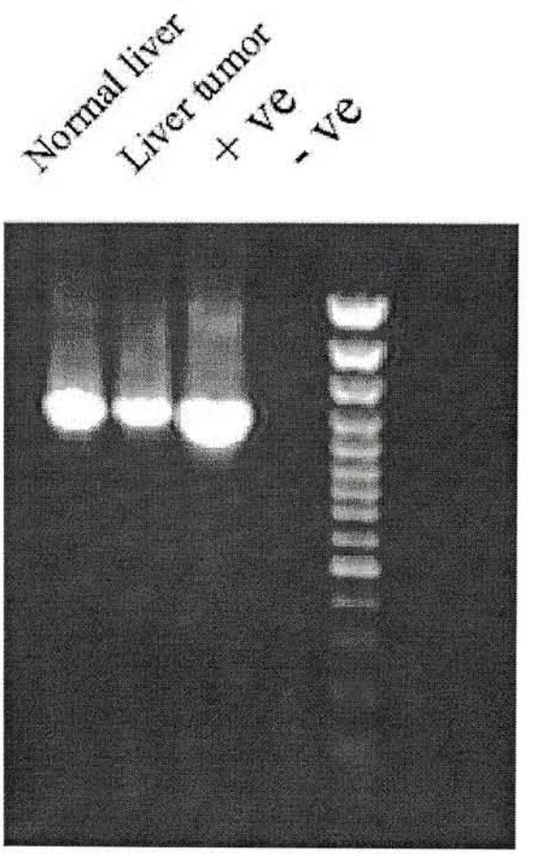


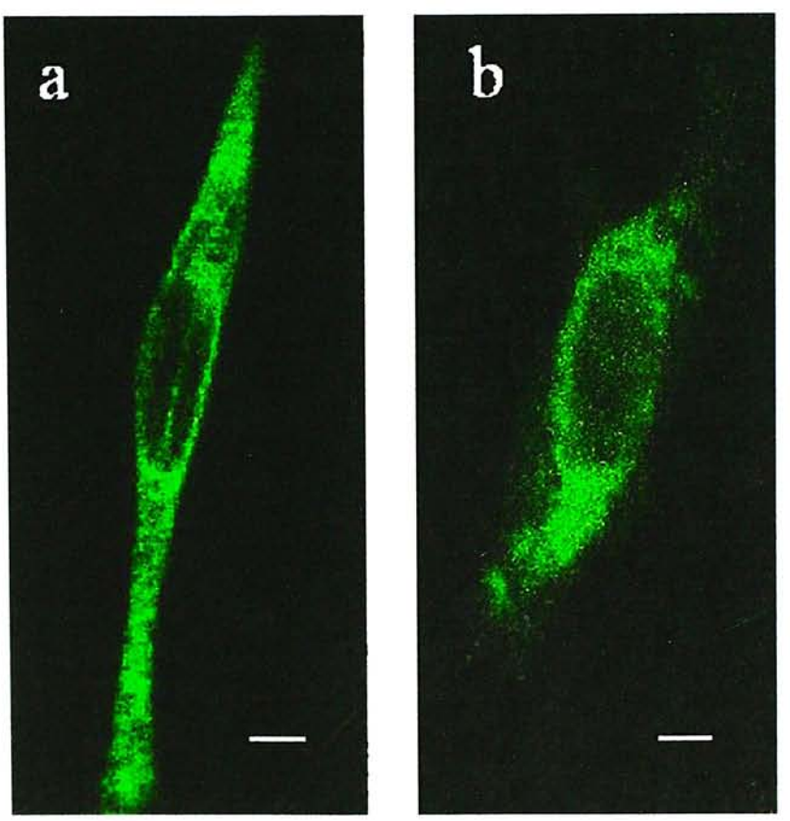
Figure 3.16 PCR detection of hOAT2 partial sequence from human normal liver and liver tumor first-strand cDNA producing a transcript size approximately at 1.3-kb. Expression from both tissues was shown in the figure.



3.6 Determination of subcellular localization of hOAT2 protein by conjugating fluorescence protein

The hOAT2 gene in frame with the green fluorescent protein sequence was transfected into C2C12 mammalian cells. Subcellular localization of these fusion proteins GFP-hOAT2 was examined under confocal microscope. Figure 3.17 shows that the fluorescence (488nm excitation, 509nm emission) was uniformly distributed in the cytosolic compartment but excluded from the nucleus and the membrane structures.

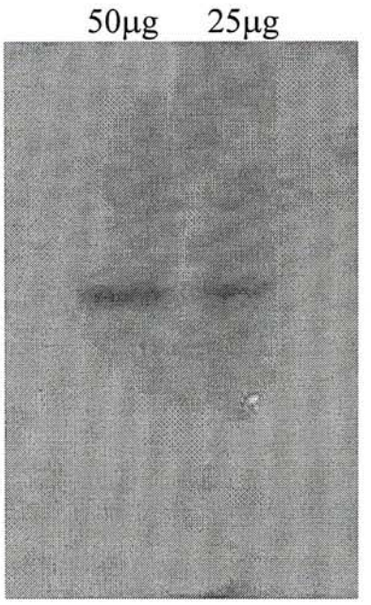
Figure 3.17 Fluorescence images of hOAT2 conjugated with GFP in **a.** C2C12 and **b.** HepG2 cells by confocal microscope. Cytosolic localization of the fusion proteins is clearly observed. Scale bar, 5μm in all fluorescence images.



3.7 Immunodetection of protein extracts from cultured cells

Polyclonal antibodies against hOAT2 were utilized for testing various protein extracts. Mammalian cells Hep3B was one of them used in testing against the anti-hOAT2 antibodies. Both lanes showed in figure 3.18 were loaded with extracted proteins from Hep3B in different amount. Discrete bands are observed in both lanes with protein size around 60kDa. Whether recognition of anti-hOAT2 antibodies to other normal tissues or cancer cells needs further investigation.

Figure 3.18 Immunohybridization using polyclonal anti-hOAT2 antibody detected a band approximately 60kDa in Hep3B total proteins (25μg / 50μg).



4 Chapter 4: Discussion

4.1 Characterization of the hepatocellular ESTs

4.1.1 Classification and frequency distribution of the 500 ESTs

Under the sequence alignment search in BLAST, one should concern whether a given alignment constitutes evidence for homology, and is not the outcome by chance alone. To lower the possibility in case of aligned sequences that are generated randomly from the DNA or protein pool, a statistical Poisson *P*-value was utilized to figure out the reliability of those alignments (Altschul and Erickson, 1985; Fitch, 1983; Lipman *et al.*, 1984). For example, a *P*-value < 0.01 can say reliably as if 100 random alignments have score inferior to the alignment of interest.

The 500 ESTs presented in this project were a combination of sequences generated from two cDNA libraries, human normal liver and human cancerous liver as shown in the Table 3.1. It seems that the number of 500 might not be sufficient for representing a significant gene expression pattern of these tissues, it could more or less, give us an insight to the distribution of ESTs. Upon non-redundant search of the 300 normal liver ESTs, 224 of them (75%) matched with known genes, 43 (14%) of them

matched with novel genes and the remaining 33 (11%) of them were regarded as mitochodrial genes. On the other side, similar pattern was likely to be observed in the 200 liver tumor ESTs, whereas 67% (134) of known genes, 23% (46) of novel genes and the last 10% (20) was grouped into the mitochondrial category. It could be seen that there are relatively low percentages of genes that were identified as novel when compared with data from the last few years. It implies that the exploration of genes is exponentially increasing in this appreciable period, with the aid of improved sequencing technique, instruments and advanced information technologies.

Identified known genes, in accordance with their main functions, were further divided into eleven subgroups as described previously in Table 3.2. The overall distribution of ESTs between these two libraries was roughly similar except two categories, transcription / translation proteins and secretory proteins. An extraordinary high abundance of transcription / translation proteins in cancer liver (31.3%) than in normal (3.1 %), might link to the rapid and frequent divisions of cancerous cells which requires active synthesis of cellular materials such as nuclear DNA and structural components. A comparatively large discrepancy existed in the expression of secretory proteins, though it was quite abnormal to see such a high

frequency in normal liver (65.2%) because such unexpected high frequency of secretory proteins is not observed in other liver EST library (Adams *et al.*, 1992). Thus note to this point, it is necessary for us to generate significant number of ESTs before coming to any conclusion.

4.1.2 The expression pattern of membrane associated proteins

This large-scale liver ESTs sequences are the cooperative works contributed by many members in our laboratory. The gene frequency distribution classification are compatible with the one performed by my colleague (C.H. Yu, M.Phil thesis). The ESTs encoding membrane-associated proteins only accommodates low percentages for normal (3.3%), tumor-normal (3.6%) and tumor (2%) EST libraries. The low abundance of membrane-associated ESTs expressing in these liver tissues might represent the characteristic physiological and regulatory functions of this organ. For example, a rather high abundance of membrane-associated proteins in human brain (15.3%) (Adams *et al.*, 1993) is likely related to its blood-brain barrier to provide limited substrates, hormones or ions across the brain, plus the huge energy consumption via Na^+K^+ATPase which transport Na^+ and K^+ across plasma membranes.

4.2 Tissue distribution and expression profiles of hOAT2

RNA blot analysis revealed that rat OAT2 protein has a strong expression in liver and relatively weak expression in kidney (Simonson *et al.*, 1994). In order to examine the expression pattern of human OAT2, Northern blot analysis was performed. Unfortunately, our multiple tissue blot did not show hOAT2 expression. Multiple bands of irrelevant transcript sizes and also weak expression were detected. This discrepancy in the Northern blot analysis may not be attributed to the specificity of the DNA probes, since probes constructed from different regions of the cDNA had been selected. We chose the 5'-coding region of hOAT2 of sizes 243 and 570 that had < 40% identities to hOAT1, hOAT3 and hOAT4. For this reason, we make use of RT-PCR (Clontech) for the expression studies. It has the advantage of high sensitivity, and it is possible to detect some rare transcripts that cannot be determined in Northern blots. As shown in figure 3.14, hOAT2 was detected at a high level in liver and at a lower level in kidney, obviously, the result was consistent to what was determined in rat.

In comparing with other human OAT members, where hOAT1 and hOAT3 were determined as kidney-specific proteins (Race *et al.*, 1999) and hOAT4 was expressed in both kidney and placenta (Hosoyamada *et al.*,

1999), a common characteristic could be seen from these expression profiles including hOAT2. Interestingly, the kidney seems to be unique for all of the four organic anion transporters, though some of them were not predominantly expressed in that particular organ. Whether the existence of common substrate(s) exists between the four members need further experimental studies.

Other types of transport proteins such as Ntcp, oatp1 and OCT were found to be located at the sinusoidal hepatocyte membrane (Kullak-Ublick, 1999) where rOAT2 is also expressed in that region (Simonson *et al.*, 1994). Nevertheless, the substrate specificities among these transporters are quite divergent. Ntcp and oatp1 were believed to be bile acid transporters, which mediates the Na^+-dependent hepatocellular uptake of conjugated bile acids such as taurocholate, and Na^+-independent uptake of taurocholate, cholate, anionic steroid conjugates (Jacquemin *et al.*, 1994).

4.3 HOAT2 in fetal development

The pattern of expression of the rat OAT2 gene in liver development had been demonstrated by Simonson *et al.* (1994). At the fetal stage, the level of expression was very low, however it dramatically increased at the

time of birth and remained active throughout adulthood (Simonson *et al.*, 1994). Amplification of human fetal multiple tissue cDNA revealed that the hOAT2 transcript was expressed in fetal liver but was absent in fetal kidney. It is different from previous results where hOAT2 was expressed in both human adult liver and kidney. In our knowledge, it is the first finding showing differential expression of hOAT2 in fetal and adult kidney. Such prior maturation of liver to kidney might suggest that there is a higher need for the removal of various organic compounds by enterohepatic circulation than in kidney (Gregus Z., 1991).

4.4 Predicting the topology of membrane proteins

Many of the transport proteins are believed, on the basis of their amino acid sequences and hydropathy plots, to have multiple membrane-spanning helical regions. Theoretical prediction algorithms have been found to be useful in detecting membrane-spanning regions from the primary structure alone. One of the most common use methods is adopted by Kyte and Doolittle (1982). The calculation is involved the mean residue hydrophobicity values obtained for consecutive 19-residue spans, where segments with hydrophobicity above a certain threshold are predicted to be

membrane-spanning.

Topology prediction of hOAT2 showed twelve transmembrane segments as described in figure 3.4 and table 3.5. In addition to a large extracellular loop (101 amino acids) on the region between the putative transmembrane domains one and two, a comparatively smaller intracellular loop with 67 amino acid residues is located between domains six and seven. It is generally believed that the recognition of extracellular substrates or anions requires the interacting hydrophilic surfaces, such as those present in the N-terminus and the extracellular loop. So far, both the hydrophilic N-terminus and C-terminus of hOAT2 are predicted to be located in the intracellular compartment, the discrete hydrophilic loop becoming the focus for the transporter in substrate recognition and even its multispecificity.

It should be noted that these assignments of topology should be considered tentative until direct structural determination, such as X-ray diffraction analysis or electron microscope scanning, has been performed.

4.5 Chromosomal mapping of human OAT2

Radiation hybrid (RH) mapping makes use of somatic cell genetic approach using a random set of irradiated human donor cells to non-

irradiated hamster cells. In the GeneBridge 4 RH mapping panel, approximately 30 % on average of the donor genomic DNA is stably retained in the hamster cells. A 3000 rad of X-ray dosage is applied in constructing the RH map with the average fragment size of 25 Mb and 1 Mb in effective resolution (Gyapay *et al.*, 1996).

The human OAT2 was localized to chromosome 6 approximately 249.63cR from the top of chromosome linkage group, which is equivalent to 6p21.2 to p21.1 with reference that $1cR_{3000}$ of X-ray breakage representing 270kb in DNA distance. Another two members, hOAT1 and hOAT3, both were mapped to 11q11.7 (Race *et al.*, 1999). From this apparent result, a closer association is found between hOAT1 and hOAT3 than either of those with hOAT2. Moreover, the hOAT2 gene is located distal to haemochromatosis (HFE) gene at the position 6p23 (Feder *et al.*, 1996) and is proximal to acidic epididymal glycoprotein-like 1 (AEGL1) gene (Hayashi *et al.*, 1996) and peroxisome proliferative activated receptor, delta (PPARD) gene (Schmidt *et al.*, 1992) which are placed at the same location 6p21.2 – p21.1. The rapid increases in discovery of gene locations by latest advances of mapping technique such as fluorescence in *situ* hybridization and radiation hybrid mapping have given us an insight on some disease related genes. Although in most cases, diseases could be multifactorial and seldom

induced directly by expression or suppression of a single gene, at least some mutational hot-spots or some active translocation sites could be deduced by gathering information from the gene locations.

4.6 Possible functions of hOAT2

4.6.1 Hepato-renal relation

Liver is a vital organ in maintaining the clearance and metabolism of endogenous and xenobiotics substances. Bile acid or bile salts synthesized from the liver are major constituents of bile and involve actively in the enterohepatic circulation. For example, Ntcp and Oatp are the two major bile salt transport systems locating at the basolateral membrane for the hepatic extraction of bile salts from portal blood plasma into bile (Hagenbuch *et al.*, 1991; Jacquemin *et al.*, 1994). In the other side, functional tests on organic anion transporters in kidney have being widely studied. Different endogenous and exogenous organic anions including metabolites and drugs have been examined to gather information about the transport of these compounds as well as their nephrotoxicities. *p*-Aminohippurate (PAH) is one of the indicator which has long been used for the estimation of the renal plasma flow (Hirsch, 1976), and is known to be mediated by multispecific transporters such as OAT1 (Hosoyamada *et al.*, 1999; Sekine *et*

al., 1997) and OAT3 (Kusuhara *et al.*, 1999; Race *et al.*, 1999). It has been extensively used as a test substrate in analyzing transport efficacy of organic anion transporters in response to various drugs and endogenous substrate treatments.

To date, even a number of organic anions transporters have been identified, it is not obvious how to deduce relationships between various organic anions or substrates carried by those particular transporters in either molecular structures or physiological properties. Salicylate, acetylsalicylate, PGE2, dicarboxylate and PAH are substrates for rat OAT2 (Sekine *et al.*, 1998). However, other transporters could also mediate those organic anions, such as PAH transported by OAT1 and OAT3 (Sekine *et al.* 1997). It is not hard to see in other kind of transporters like Oatp1, Oatp2, Oatp3 and Ntcp, they all have a common substrate, thyroxine (Kullak-Ublick, 1999). It seems that most organic anion transporters may specify a group of substrates, but overlapping of these anion groups might occur with even more than one kind of substrate involved. ROAT2 is structurally similar to renal transporter rOAT1 as well as its substrate selectivity, however, the transport efficacy is different. As described, the PAH transport by rOAT2 occurred at a much lower level than that by rOAT1 (Sekine *et al.*, 1997) but the mediation of salicylate by rOAT2 was much higher (Sekine

et al., 1998). To this point, it had been speculated that the tissue distribution of anions was different in the body, with evidence in previous finding showing that PAH is excreted mainly into the urine and rarely into the bile (Hirsch, 1976).

Investigation on the hepato-renal relationship can provide us information on the transport mechanisms and the substrate specificity among the organic anion transporters.

4.6.2 Substrate diversity

Multispecificity is the common characteristic in the majority of organic anion transporters. In fact, these transporters are said to be multispecific rather than non-specific because they mediate many endogenous and exogenous compounds, but not any lipophilic or hydrophilic compounds. The four human OAT isoforms, hOAT1, hOAT2, hOAT3 and hOAT4 have been shown to be expressed in kidney, though in different expression levels (Cha *et al.*, 2000; Hosoyamada *et al.*, 1999; Race *et al.*, 1999; Sekine *et al.*, 1998). To maintain the clearance of a wide variety of nephrotoxic compounds and metabolites from the body, the OAT family plays a key role in transporting these anions in kidney as well as in the metabolic center, liver. Moreover, such diverse substrate specificity property may serve another purpose. These multispecific organic anion

transporters have been deduced to own a characteristic substrate group, but it is believed there are still a number of substrates that have not been elucidated yet. Once the body has to deal with facing the attack of some kind of undefined toxic compounds that may generated exogenously or endogenously, these transporters could have the ability to deal with the toxin, followed by detoxification or excretion, and prevent accumulation in the body.

It has been identified that some clinically important drugs are transported by or have inhibitory effect on the OATs. Uptake of drug such as methotrexate was found to be mediated by rOAT2 (Sekine *et al.*, 1998); furosemide (a loop diuretic), probenecid (a uricosuric drug) and indomethacin (a nonsteroidal anti-inflammatory drug) are regarded as inhibitors in hOAT1-mediated PAH uptake (Hosoyamada *et al.*, 1999). Furthermore, antiviral drugs like cidofovir and adefovir are novel substrates for both hOAT1 and rOAT1, and accumulation of these drugs probably associated with nephrotoxicity (Cihlar *et al.*, 1999).

4.7 Fluorescence detection for subcellular localization

Immunohistochemical studies on OAT1 proteins revealed that it is expressed exclusively on the basolateral membrane of S2 segments of proximal tubules in rat kidney, and none was found in the cytoplasmic vesicles (Tojo *et al.*, 1999).

Accordingly, they suggested that organic anions might not be absorbed via an endocytotic pathway. Immunostaining using monoclonal antibody 2C5 detected rOAT2 on sinusoidal domains of hepatocytes but absence on the bile canalicular membrane (Simonson *et al.*, 1994). Nevertheless, our fluorescence detection results showed the GFP tagged hOAT2 fusion proteins were localized evenly in the cytosolic area of mammalian cells, C2C12. The result was consistent when using hepatocarcinoma cell line, HepG2. Although hOAT2 is a putative transmembrane protein in regard to the preliminary analysis of its amino acid structure and hydropathicity, the transport mechanism or anions carrying pathway is still incompletely understood. Our results do not rule out the possibility that the GFP tag affected the localization of OAT2 adversely, and further work needs to be done using GFP tagged to a C-terminus of the protein.

4.8 Conclusion

We have identified and characterized hOAT2 from the screening of human liver cDNA library. HOAT2 is believed to be the human homolog of rat OAT2 (NLT) since it shows 88% nucleotide sequence identity, and becomes the second member of multispecific organic anion transporter family. Primary structural analysis revealed that it possesses twelve putative transmembrane domains and a

large extracellular loop between transmembrane segment one and two, plus a possible *N*-glycosylatoin site within the loop.

RT-PCR on human adult tissues have demonstrated that hOAT2 has a high level of expression in liver and relatively weak expression in kidney, while expression was only observed in liver when analyzed on human fetal tissues. The expression pattern of adult human tissues is consistent with that of rat. Chromosomal localization of hOAT2 has been deduced to chromosome 6p21.2 – > p21.1 using radiation hybrid mapping.

The subcellular localization of hOAT2 was analyzed by using green fluorescence technique. It is surprising that hOAT2 was found to be localized in the cytoplasm, even in either liver or non-liver cell lines. Further studies are necessary to confirm this finding.

Expression of hOAT2 was detected in liver cancer, though not quantitatively different from that of normal liver. We have raised an antibody against hOAT2. Western blotting using anti-hOAT2 antibodies deduced a band around 60 kDa on protein extract from Hep3B cells which is equivalent to the size of hOAT2 protein.

4.9 Further aspects

Up till now, functional studies on hOAT2 are still at an initial stage. On the rat model, we know alpha-ketoglutarate and PAH are the substrates for rOAT2. In addition, hOAT2 was found to be expressed in liver cancer tissue. Transfection of hOAT2 into HepG2 cells could help us to investigate the transport efficacy of this transporter, and this strategy had been employed by Kullak-Ublick *et al.* in analyzing bile acid transport by Oatp (1996). The uses of radiolabelled substrates and antisense oligonucleotides from different regions are helpful in chasing the amount of the substrate uptake and determining the binding sequence of that protein respectively. Moreover, drugs and other substrates could be added to the stably transected cells to test whether they are inhibitory towards the hOAT2.

References

Adams, M. D., Dubnick, M., Kerlavage, A. R., Moreno, R., Kelley, J. M., Utterback, T. R., Nagle, J. W., Fields, C., and Venter, J. C. (1992). Sequence identification of 2,375 human brain genes [see comments]. Nature ***355***, 632-4.

Adams, M. D., Kelley, J. M., Gocayne, J. D., Dubnick, M., Polymeropoulos, M. H., Xiao, H., Merril, C. R., Wu, A., Olde, B., Moreno, R. F., and et al. (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. Science ***252***, 1651-6.

Adams, M. D., Kerlavage, A. R., Fields, C., and Venter, J. C. (1993). 3,400 new expressed sequence tags identify diversity of transcripts in human brain. Nat Genet ***4***, 256-67.

Altschul, S. F., Boguski, M. S., Gish, W., and Wootton, J. C. (1994). Issues in searching molecular sequence databases. Nature Genetics ***6***, 119-129.

Altschul, S. F., and Erickson, B. W. (1985). Significance of nucleotide sequence alignments: a method for random sequence permutation that preserves dinucleotide and codon usage. Mol Biol Evol ***2***, 526-38.

Bairoch, A., and Murzin, A. G. (1997). Sequences and topology: Predicting evolution. Curr Opin Struct Biol ***7***, 367-8.

Bause, E. (1983). Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes. Biochem J ***209***, 331-6.

Bolder, U., Ton-Nu, H. T., Schteingart, C. D., Frick, E., and Hofmann, A. F. (1997). Hepatocyte transport of bile acids and organic anions in endotoxemic rats: impaired uptake and secretion [see comments]. Gastroenterology ***112***, 214-25.

Bossard, R., Stieger, B., O'Neill, B., Fricker, G., and Meier, P. J. (1993). Ethinylestradiol treatment induces multiple canalicular membrane transport alterations in rat liver. J Clin Invest ***91***, 2714-20.

Cha, S. H., Sekine, T., Kusuhara, H., Yu, E., Kim, J. Y., Kim, D. K., Sugiyama, Y., Kanai, Y., and Endou, H. (2000). Molecular cloning and characterization of

multispecific organic anion transporter 4 expressed in the placenta. J Biol Chem ***275***, 4507-12.

Cihlar, T., Lin, D. C., Pritchard, J. B., Fuller, M. D., Mendel, D. B., and Sweet, D. H. (1999). The antiviral nucleotide analogs cidofovir and adefovir are novel substrates for human and rat renal organic anion transporter 1. Mol Pharmacol ***56***, 570-80.

Endou, H. (1998). Recent advances in molecular mechanisms of nephrotoxicity. Toxicol Lett ***102-103***, 29-33.

Feder, J. N., Gnirke, A., Thomas, W., Tsuchihashi, Z., Ruddy, D. A., Basava, A., Dormishian, F., Domingo, R., Jr., Ellis, M. C., Fullan, A., Hinton, L. M., Jones, N. L., Kimmel, B. E., Kronmal, G. S., Lauer, P., Lee, V. K., Loeb, D. B., Mapa, F. A., McClelland, E., Meyer, N. C., Mintier, G. A., Moeller, N., Moore, T., Morikang, E., Wolff, R. K., and et al. (1996). A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis [see comments]. Nat Genet ***13***, 399-408.

Fitch, W. M. (1983). Random sequences. J Mol Biol ***163***, 171-6.

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., and et al. (1995). Whole-genome random sequencing and assembly of Haemophilus influenzae Rd [see comments]. Science ***269***, 496-512.

Frimmer, M., and Ziegler, K. (1988). The transport of bile acids in liver cells. Biochim Biophys Acta ***947***, 75-99.

Gartung, C., Ananthanarayanan, M., Rahman, M. A., Schuele, S., Nundy, S., Soroka, C. J., Stolz, A., Suchy, F. J., and Boyer, J. L. (1996). Down-regulation of expression and function of the rat liver Na^+/bile acid cotransporter in extrahepatic cholestasis. Gastroenterology ***110***, 199-209.

Gerloff, T., Geier, A., Stieger, B., Hagenbuch, B., Meier, P. J., Matern, S., and Gartung, C. (1999). Differential expression of basolateral and canalicular organic anion transporters during regeneration of rat liver. Gastroenterology ***117***, 1408-15.

Goldinger, J. M., Erasmus, B. D., Song, Y. K., Koschier, F. J., and Hong, S. K. (1980). Effects of SCN^- and $NO3^-$ on organic anion transport in rabbit kidney cortical slices. Biochim Biophys Acta ***598***, 357-65.

Green, R. M., Beier, D., and Gollan, J. L. (1996). Regulation of hepatocyte bile salt transporters by endotoxin and inflammatory cytokines in rodents. Gastroenterology ***111***, 193-8.

Gyapay, G., Schmitt, K., Fizames, C., Jones, H., Vega-Czarny, N., Spillett, D., Muselet, D., Prud'Homme, J. F., Dib, C., Auffray, C., Morissette, J., Weissenbach, J., and Goodfellow, P. N. (1996). A radiation hybrid map of the human genome. Hum Mol Genet ***5***, 339-46.

Gyurasics, A., Varga, F., and Gregus, Z. (1991). Glutathione-dependent biliary excretion of arsenic. Biochem Pharmacol ***42***, 465-8.

Hagenbuch, B., and Meier, P. J. (1994). Molecular cloning, chromosomal localization, and functional characterization of a human liver Na^+/bile acid cotransporter. J Clin Invest ***93***, 1326-31.

Hagenbuch, B., Stieger, B., Foguet, M., Lubbert, H., and Meier, P. J. (1991). Functional expression cloning and characterization of the hepatocyte Na^+/bile acid cotransport system. Proc Natl Acad Sci U S A ***88***, 10629-33.

Handler, J. S., Green, N., and Steele, R. E. (1989). Cultures as epithelial models: porous-bottom culture dishes for studying transport and differentiation. Methods Enzymol ***171***, 736-44.

Hayashi, M., Fujimoto, S., Takano, H., Ushiki, T., Abe, K., Ishikura, H., Yoshida, M. C., Kirchhoff, C., Ishibashi, T., and Kasahara, M. (1996). Characterization of a human glycoprotein with a potential role in sperm-egg fusion: cDNA cloning, immunohistochemical localization, and chromosomal assignment of the gene (AEGL1). Genomics ***32***, 367-74.

Hirsch, G. H. (1976). Differential effects of nephrotoxic agents on renal transport and metabolism by use of in vitro techniques. Environ Health Perspect ***15***, 89-99.

Hosoyamada, M., Sekine, T., Kanai, Y., and Endou, H. (1999). Molecular cloning

and functional expression of a multispecific organic anion transporter from human kidney. Am J Physiol ***276***, F122-8.

Inui, K., Takano, M., Okano, T., and Hori, R. (1986). Role of chloride on carrier-mediated transport of p-aminohippurate in rat renal basolateral membrane vesicles. Biochim Biophys Acta ***855***, 425-8.

Jacquemin, E., Hagenbuch, B., Stieger, B., Wolkoff, A. W., and Meier, P. J. (1994). Expression cloning of a rat liver Na(+)-independent organic anion transporter. Proc Natl Acad Sci U S A ***91***, 133-7.

Jung, K. Y., and Endou, H. (1989). Nephrotoxicity assessment by measuring cellular ATP content. II. Intranephron site of ochratoxin A nephrotoxicity. Toxicol Appl Pharmacol ***100***, 383-90.

Khan, A. S., Wilcox, A. S., Polymeropoulos, M. H., Hopkins, J. A., Stevens, T. J., Robinson, M., Orpana, A. K., and Sikela, J. M. (1992). Single pass sequencing and physical and genetic mapping of human brain cDNAs [see comments]. Nat Genet ***2***, 180-5.

Kornfeld, R., and Kornfeld, S. (1985). Assembly of asparagine-linked oligosaccharides. Annu Rev Biochem ***54***, 631-64.

Kullak-Ublick, G. A. (1999). Regulation of organic anion and drug transporters of the sinusoidal membrane. J Hepatol ***31***, 563-73.

Kusuhara, H., Sekine, T., Utsunomiya-Tate, N., Tsuda, M., Kojima, R., Cha, S. H., Sugiyama, Y., Kanai, Y., and Endou, H. (1999). Molecular cloning and characterization of a new multispecific organic anion transporter from rat brain. J Biol Chem ***274***, 13675-80.

Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. J Mol Biol ***157***, 105-32.

Lee, S. M. Y., Tsui, S. K. W., Chan, K. K., Kotaka, M., Li, H. Y., Chim, S. S., Waye, M. M. Y., Fung, K. P., and Lee, C. Y. (1998). Chromosomal mapping of a skeletal muscle specific LIM-only protein FHL3 to the distal end of the short arm of human chromosome 1. Somat Cell Mol Genet ***24***, 197-202.

Lipman, D. J., Wilbur, W. J., Smith, T. F., and Waterman, M. S. (1984). On the statistical significance of nucleic acid similarities. Nucleic Acids Res ***12***, 215-26.

Meier, P. J. (1995). Molecular mechanisms of hepatic bile salt transport from sinusoidal blood into bile. Am J Physiol ***269***, G801-12.

Okubo, K., Hori, N., Matoba, R., Niiyama, T., Fukushima, A., Kojima, Y., and Matsubara, K. (1992). Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression [see comments]. Nat Genet ***2***, 173-9.

Olson, J. M., Boehnke, M., Neiswanger, K., Roche, A. F., and Siervogel, R. M. (1989). Alternative genetic models for the inheritance of the phenylthiocarbamide taste deficiency. Genet Epidemiol ***6***, 423-34.

Petzinger, E., Follmann, W., Blumrich, M., Walther, P., Hentschel, J., Bette, P., Maurice, M., and Feldmann, G. (1994). Immortalization of rat hepatocytes by fusion with hepatoma cells. I. Cloning of a hepatocytoma cell line with bile canaliculi. Eur J Cell Biol ***64***, 328-38.

Pfeffer, S. R., and Rothman, J. E. (1987). Biosynthetic protein transport and sorting by the endoplasmic reticulum and Golgi. Annu Rev Biochem ***56***, 829-52.

Race, J. E., Grassl, S. M., Williams, W. J., and Holtzman, E. J. (1999). Molecular cloning and characterization of two novel human renal organic anion transporters (hOAT1 and hOAT3). Biochem Biophys Res Commun ***255***, 508-14.

Rodriguez-Boulan, E., and Salas, P. J. (1989). External and internal signals for epithelial cell surface polarization. Annu Rev Physiol ***51***, 741-54.

Schmidt, A., Endo, N., Rutledge, S. J., Vogel, R., Shinar, D., and Rodan, G. A. (1992). Identification of a new member of the steroid hormone receptor superfamily that is activated by a peroxisome proliferator and fatty acids. Mol Endocrinol ***6***, 1634-41.

Schmitt, C., and Burckhardt, G. (1993). Modulation by anions of p-aminohippurate transport in bovine renal basolateral membrane vesicles. Pflugers Arch ***425***, 241-7.

Sekine, T., Cha, S. H., Tsuda, M., Apiwattanakul, N., Nakajima, N., Kanai, Y., and Endou, H. (1998). Identification of multispecific organic anion transporter 2 expressed predominantly in the liver. FEBS Lett ***429***, 179-82.

Sekine, T., Watanabe, N., Hosoyamada, M., Kanai, Y., and Endou, H. (1997). Expression cloning and characterization of a novel multispecific organic anion transporter. J Biol Chem ***272***, 18526-9.

Shneider, B. L., Fox, V. L., Schwarz, K. B., Watson, C. L., Ananthanarayanan, M., Thevananther, S., Christie, D. M., Hardikar, W., Setchell, K. D., Mieli-Vergani, G., Suchy, F. J., and Mowat, A. P. (1997). Hepatic basolateral sodium-dependent-bile acid transporter expression in two unusual cases of hypercholanemia and in extrahepatic biliary atresia. Hepatology ***25***, 1176-83.

Simon, F. R., Fortune, J., Iwahashi, M., Gartung, C., Wolkoff, A., and Sutherland, E. (1996). Ethinyl estradiol cholestasis involves alterations in expression of liver sinusoidal transporters. Am J Physiol ***271***, G1043-52.

Simonson, G. D., Vincent, A. C., Roberg, K. J., Huang, Y., and Iwanij, V. (1994). Molecular cloning and characterization of a novel liver-specific transport protein. J Cell Sci ***107***, 1065-72.

Tatusova, T. A., and Madden, T. L. (1999). BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences [published erratum appears in FEMS Microbiol Lett 1999 Aug 1;177(1):187-8]. FEMS Microbiol Lett ***174***, 247-50.

Tojo, A., Sekine, T., Nakajima, N., Hosoyamada, M., Kanai, Y., Kimura, K., and Endou, H. (1999). Immunohistochemical localization of multispecific renal organic anion transporter 1 in rat kidney. J Am Soc Nephrol ***10***, 464-71.

Tusnady, G. E., and Simon, I. (1998). Principles governing amino acid composition of integral membrane proteins: application to topology prediction. J Mol Biol ***283***, 489-506.

Wilcox, A. S., Khan, A. S., Hopkins, J. A., and Sikela, J. M. (1991). Use of 3'

untranslated sequences of human cDNAs for rapid chromosome assignment and conversion to STSs: implications for an expression map of the genome. Nucleic Acids Res ***19***, 1837-43.

Fluorescent vector pEGFP_C1

pEGFP-C1 Vector Information PT3028-5

GenBank Accession #: U55763 Catalog #6084-1